

Home ovulation test use and stress during subfertility evaluation: Subarm of a randomized controlled trial

Women's Health
Volume 15: 1–12
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1745506519838363
journals.sagepub.com/home/whe



Sarah Weddell¹, Georgina L Jones², Sheila Duffy³,
Cameron Hogg⁴ , Sarah Johnson¹ and William Ledger⁵

Abstract

Objectives: A prospective, randomized controlled trial in women seeking to conceive examined the impact of using ovulation tests on self-reported levels of stress, psychological well-being, and quality of life in women with unexplained infertility.

Method: The test group used a home ovulation test to detect the day of ovulation, whereas the control group were provided with a predicted day of ovulation based on the average length of menstrual cycle reported during study recruitment. Volunteers collected their first morning urine samples to evaluate biochemical levels of stress (urinary cortisol and estrone-3-glucuronide) and completed questionnaires over two complete menstrual cycles.

Results: Overall, the use of digital ovulation tests by sub-fertile women under medical care had negligible negative effects and no detectable positive benefit on psychological well-being, according to multiple measurements of stress by questionnaire and biochemical markers. No significant differences were found between groups for all stress measures at the various study time points, except in relation to “couple concordance” where the test group scored much higher than the control group (mean difference at end of study was 21.25 (95% confidence interval: 9.25, 33.25; $P=0.0015$)). The maximum difference in log cortisol: creatinine ratio between the test and control groups was -0.28 (95% confidence interval: $-0.69, 0.13$).

Conclusions: These results do not support propositions that using digital ovulation tests can cause stress in women trying to conceive.

Keywords

cortisol, infertility, ovulation tests, questionnaire, stress, timed intercourse

Date received: 14 May 2018; revised: 27 July 2018; accepted: 22 February 2019

Introduction

It has been estimated that in the United Kingdom, one in six couples who try to conceive experience some form of infertility, resulting in over 52,000 women per year seeking specialist fertility treatment.¹ Failure to become pregnant and coping with infertility have often been reported as major life stressors.^{2–5}

As conception most often occurs after unprotected sexual intercourse during a 6-day fertile window (which ends on the day of ovulation), incorrect timing of intercourse is thought to be a common, but easily correctable, cause of difficulties in conceiving.^{6,7} Ovulation tests detect the surge in levels of luteinizing hormone (LH) that precedes ovulation and can, therefore, prospectively identify a woman's most fertile phase.

Social trends in developed countries, such as deferred parenthood and poor awareness of the phases of one's own

¹Scientific and Medical Affairs, SPD Swiss Precision Diagnostics Development Company Limited, Bedford, UK

²Department of Psychology, School of Social Sciences, Leeds Beckett University, Leeds, UK

³Department of Obstetrics and Gynaecology, Royal Hallamshire Hospital, Sheffield, UK

⁴Statistics and Data Management, SPD Development Company Limited, Bedford, UK

⁵School of Women's and Children's Health, University of New South Wales, Sydney, NSW, Australia

Corresponding author:

Sarah Johnson, Scientific and Medical Affairs, SPD Swiss Precision Diagnostics Development Company Limited, Bedford MK44 3UP, UK.
Email: sarah.johnson@spdspark.com



menstrual cycle, often mean that women are unable to become pregnant as fast as they would like.⁸ Increased ability to identify the fertile phase could help women to achieve pregnancy sooner.⁹ Among women trying to conceive, ovulation tests are a simple, effective, and popular method to maximize the chances of a natural conception.^{6,8} These devices may also be useful in a clinical setting as a convenient way to accurately schedule procedures such as egg collection in a “natural” assisted reproduction cycle and embryo transfer in a frozen–thawed cycle.¹⁰

Despite these potential benefits, timing sexual intercourse using ovulation tests has been reported to be associated with emotional distress.¹¹ Between 2004 and 2013, the UK National Institute for Health and Care Excellence (NICE) guidelines indicated that timing sexual intercourse so that it coincides with the period of maximum fertility, assessed by ovulation testing or other methods, causes stress.¹²

Despite the above-mentioned association between ovulation tests and stress, to date, the only controlled study to examine this showed that there was no negative impact on women's psychological well-being.¹³ In addition to this, the use of a digital ovulation test appeared to increase the likelihood of achieving pregnancy in a population of women trying to conceive in a non-medical setting who were at the early stages of their reproductive journey. These results are not applicable to women who had previous problems with conception or were diagnosed with infertility. To date, our knowledge remains limited with regard to the association between ovulation test use and stress levels in women undergoing a psychologically challenging process of infertility diagnosis or treatment who still wish to try to conceive naturally prior to commencing treatment.

The present study was conducted in a subgroup of patients included in the above study, with the objective of assessing whether the use of home ovulation tests affects the levels of self-reported stress, psychological well-being, and quality of life in women recently diagnosed with “unexplained” infertility.

Methods

Study design and recruitment

This prospective, randomized controlled trial involved women referred for infertility treatment at the Fertility Clinic, Jessop Wing, Royal Hallamshire Hospital in Sheffield (UK), aiming to examine whether stress levels in women seeking to conceive are affected by use of home ovulation tests. Women were either randomized to the test group or the control group. Test group volunteers were given the Clearblue Digital Ovulation Test (SPD Swiss Precision Diagnostics Development Company Limited, Bedford, UK) for the duration of the study, and asked to begin testing on day 6 of their menstrual cycle and continue daily testing until they reached surge in levels of LH.

Control group volunteers were specifically asked not to identify their day of ovulation using methods such as ovulation prediction kits or basal body temperature measurement. Both groups were given written advice concerning timing of sexual intercourse to maximize chances of conception—that is, sexual intercourse every 2–3 days for the duration of the menstrual cycle. It was the volunteer's choice as to whether or not they followed these guidelines. The study was designed to last two complete menstrual cycles, with levels of stress assessed throughout the study as detailed in the study protocol section.

The study was conducted in a subgroup of patients from a larger study which involved a greater population of women trying to conceive outside of a medical setting and was approved separately by Sheffield NHS Research Ethics Committee (UK; 09/H1308/134). All experiments and procedures were performed in accordance with relevant guidelines and regulations. The study was registered in March 2010 under the clinical trial registration number NCT01084304.

Women were recruited into the study over a period of 2 years, between February 2010 and March 2012. Participants were accepted if they fulfilled the following criteria: new attendees at the fertility clinic; 18 years old or older; regular menstrual cycles between 25 and 35 days; and no diagnosis of bilateral or unilateral tubal blockage, documented anovulation, or severe male-factor infertility (with less than 5 million motile sperm per ejaculate). Those eligible to participate were given a written invitation—including a patient information leaflet—before attending the fertility clinic and were asked if they were interested in participating at the time of consultation with the fertility specialist. Those who expressed an interest met the research nurse immediately after their consultation to complete screening and written informed consent. Women who had previously used ovulation tests were not excluded from participating in the study.

Study population, sample size, and randomization

As the variability of the stress outcome measures was not known at the time of study initiation and recruitment of this subgroup took place in parallel to the main study, a pragmatic decision was made to recruit a maximum of 100 patients. However, owing to slow recruitment levels from the single tertiary referral center, over a period of approximately 2 years, the study recruitment was closed after admission of 50 patients. This was deemed to be a sufficient sample size given that a previous study by Kopitzke et al.¹¹ demonstrated a statistically significant difference in stress levels of women trying to conceive based on a sample size of 26 women within one of the study groups.

Recruited patients were randomized equally into either the test group or the control group through block randomization.

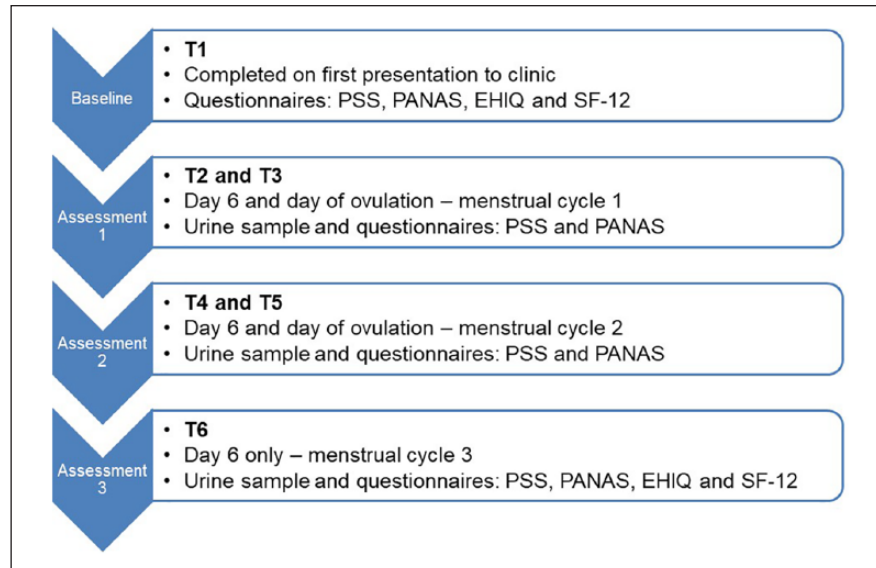


Figure 1. Assessment schedule of study protocol.

EHIQ: Emotional Health in Infertility Questionnaire; PANAS: Positive and Negative Affect Schedule; PSS: Perceived Stress Scale; SF-12v2: Short Form-12 Health Survey version 2; T: time point. Day 6 of menstrual cycle relative to the first day of menstrual bleeding.

Randomization schedules were generated using the Stata software (Statacorp Statistical Software; 2009) by a statistician who was not directly involved in the recruitment of the study participants.

Study protocol

On admission to the study, a medical history of each volunteer was collected, including demographic information, obstetric and gynecological history, and information regarding general health. Once recruited, volunteers were randomized by the study nurse according to the schedule, were given the required materials, and were asked to begin on their next menstrual cycle. For some volunteers, this resulted in pregnancy being achieved before the start of the first menstrual cycle. These were known as “pre-cycle 1 pregnancies.”

Volunteers were asked to collect their first morning urine samples and to complete a range of questionnaires at specific time points over two complete menstrual cycles according to the assessment schedule presented in Figure 1. Levels of stress were measured both biochemically and by questionnaire. A final measurement of each questionnaire and biochemical marker was taken in the third cycle to ensure that all participants completed two full cycles and to ensure that the pregnancy rate in the second menstrual cycle could be captured.

The test group used the home ovulation test daily from day 6 of their menstrual cycle until the day of ovulation (day after surge in levels of LH) was detected, whereas the control group were provided with a predicted day of ovulation based on average menstrual cycle length reported

upon recruitment to the study. This was used to calculate the most likely day of surge in levels of LH and thus day of ovulation.

Analytical procedures

Questionnaires and biochemical markers were used to assess the stress of participants in this study, as previously described in detail by Tiplady et al.¹³ The questionnaires of choice included the Perceived Stress Scale (PSS),¹⁴ the Positive and Negative Affect Schedule (PANAS),¹⁵ and the Short Form-12 Health Survey version 2 (SF-12v2).¹⁶ Total stress (as determined using the PSS)¹⁴ and both total positive affect and total negative affect (as determined by the PANAS questionnaire)¹⁵ were evaluated in test group volunteers and compared with those from control group participants at all time points (T1–T6). The SF-12v2 questionnaire scores were analyzed at the beginning (T1) and at the end (T6) of the study.

The stress biochemical marker of choice was urinary cortisol, which is a well-established stress indicator.¹⁷ In addition, urinary estrone-3-glucuronide (E3G) was analyzed as a marker of estrogen status, which was previously linked to depression and anxiety.¹⁸ Creatinine analysis was also carried out to enable volume correction of these biomarkers.

In addition, the Emotional Health in Infertility Questionnaire (EHIQ) was used in this patient population as a secondary outcome. The EHIQ is a 40-item instrument that was developed to measure emotional strain in individuals having problems trying to conceive. This tool is applicable to couples undergoing *in vitro* fertilization

treatment as it covers topics including personal strain (partner relationship strain, social support, confidence in treatment, guilt and blame, financial strain, and need for privacy). The EHIQ questionnaire scores were analyzed at the beginning (T1) and at the end (T6) of the study.

The participants' first morning urine samples were collected using sample collection pots containing sodium azide as a preservative. Upon receipt of samples at the study site, they were aliquoted and stored at a temperature of -80°C until biochemical analysis was conducted using fully validated immunoassay systems, as described by Tiplady et al.¹³

Statistical analysis

The standardized psychological questionnaires (PSS, PANAS, EHIQ, and SF-12v2) were coded following the scoring systems for each instrument and summarized using descriptive statistics. Mean questionnaire/scale scores for the volunteers were given for each time point. Summary measures were used to examine the profile of stress over time between the baselines and remaining time point assessments;¹⁹ the median was used in this instance, as scores compared with baseline results were not normally distributed. Log transformation of the biochemical marker to creatinine ratio using the natural logarithm was conducted to stabilize variance. Demographic data captured at recruitment were also analyzed to examine the influence of these covariates on total stress levels, including age, time spent trying to conceive, and alcohol consumption, among others.

For all tests with 95% confidence intervals (CIs) reported (PSS, PANAS, SF-12v2, log (cortisol/creatinine) and log (E3G/creatinine), a normal approximation using a two sample t-test with unequal variance was used. This involved using the Satterthwaite approximation which does not assume that variances are equal (SAS Stat Package; SAS Institute Inc., Cary, NC, USA).

As sample size was reduced due to difficulty in recruitment, a retrospective analysis of the effect sizes detectable with 80% and 90% power for the final sample size was conducted. This indicated that the study would have been able to detect differences of clinical significance, but differences of smaller magnitude would not be resolved. For example, a difference of 0.5 urinary cortisol/creatinine would have been detected with 90% power, or 0.43 with 80% power. Considering the questionnaires, with 90% power, a difference in score of 9.26 would have been detected for PSS (score of 8 with 80% power), and for PANAS, a score of 8.8 for the negative and 8.6 for the positive would have been detected with 90% power (7.6 negative, 7.4 positive with 80% power). In practical terms, this would equate to nine questions consistently being marked as 1 unit different on a Likert-type scale or two questions being answered diametrically opposite from

each other. Thus, differences of clinical importance would have been resolved, but not those of smaller effect size which would not be of impact on user well-being.

Data availability

The data that support the findings of this study are available on request from the corresponding author (S.J.). The data are not publicly available, as they contain information that could compromise patient privacy and/or consent.

Results

Study population

A total of 50 volunteers participated in the study. The mean age of volunteers was 32.48 years (range: 23–42 years); the mean self-reported menstrual cycle length was 28.7 days (range: 25–35 days). On average, the study volunteers tried to conceive for 22.8 months (range: 4–60 months). A summary of the demographic variables by randomization group is provided in Table 1.

All 50 volunteers were randomized upon recruitment (Figure 2); three of them (6%) did not begin the study as they became pregnant before its initiation ("pre-cycle 1 pregnancy"), and six (12%) women were either lost to follow-up or withdrew before the study started. Of the remaining 41 volunteers, 27 (64.3%) participants completed the study without achieving a pregnancy, and three (7.1%) women achieved a pregnancy at the end of the second menstrual cycle of the study. The remaining 11 (26.2%) volunteers were lost to follow-up or withdrew from the study for various reasons when the study was ongoing. As only three volunteers of the unexplained-infertility cohort became pregnant at the end of the study, comparison of conception rates between groups was not possible.

Questionnaire-based measurements of stress levels

PSS. At each time point, including baseline, the average total stress score was higher in the test group than in the control group with a maximum difference of 2.75 points (95% CI: $-2.48, 7.98$; $P=0.2914$). However, no significant difference in PSS scores between participants who used the ovulation test and those who did not was observed at any time point during the study (Figure 3(a); Supplementary Table 1).

Throughout the study, the median difference in PSS scores for each group was lower than 1 point when compared with baseline measurements, and at all time points, the 5th percentile and 95th percentile of difference contained the value zero which suggests that there is not enough evidence to show a significant difference. On

Table 1. Demographic variables reported at admission by randomization group.

| | Test group (n=25) | | | Control group (n=25) | | |
|------------------------------|-------------------|---------------|-------------|----------------------|---------------|-------------|
| | Mean | Median (SD) | Range | Mean | Median (SD) | Range |
| Age, years | 33.08 | 35 | 27–42 | 31.88 | 33 | 23–39 |
| Total previous pregnancies | 0.4 | 0 | 0–3 | 0.26 | 0 | 0–2 |
| Total live births | 0.28 | 0 | 0–1 | 0.13 | 0 | 0–1 |
| Total miscarriages | 0.06 | 0 | 0–1 | 0.27 | 0 | 0–2 |
| Months trying to conceive | 24.36 | 24 | 4–48 | 21.20 | 18 | 6–60 |
| Menstrual cycle length, days | 28.52 | 28 | 25–35 | 28.96 | 28 | 25–35 |
| Height, m | 1.65 | 1.65 (0.06) | 1.5–1.73 | 1.64 | 1.63 (0.06) | 1.52–1.76 |
| Weight, kg | 67.57 | 65.00 (12.20) | 51.7–95.3 | 66.02 | 64.00 (11.51) | 50.8–96.5 |
| BMI, kg/m ² | 25.04 | 23.39 | 18.32–36.44 | 24.65 | 24.27 | 19.19–38.17 |
| Alcohol, units/week | 3.12 | 2 | 0–12 | 4.12 | 2 | 0–14 |
| Exercise, h/week | 2.18 | 1 | 0–8 | 5.60 | 2 | 0–48 |
| Pre-cycle l period, days | 14.04 | 12 | 1–43 | 14.19 | 14.5 | 2–33 |
| Smoking history, n (%) | | | | | | |
| Yes | 0 (0) | | | 1 (4) | | |
| Ex | 1 (4) | | | 2 (8) | | |
| No | 24 (96) | | | 22 (88) | | |
| Previous ovulation test use | | | | | | |
| Yes | 17 (68) | | | 17 (68) | | |
| No | 8 (32) | | | 8 (32) | | |

BMI, body mass index; SD, standard deviation.
Numbers in brackets correspond to SDs.

average throughout the study, both test and control groups had an increase in PSS score compared with baseline levels, which at the end of the study was 1 point and 2 points for the test and control groups, respectively (Figure 3(b); Supplementary Table 2).

Positive- and negative-affect schedule. The positive-affect scores at each time point were lower in the test group, with the exception of T6. There was a significant difference at T2, in which participants using the ovulation test scored on average 4.48 points (95% CI, 0.40, 8.56; $P=0.0324$) less than control group participants (Figure 4(a); Supplementary Table 1). After T2, no significant difference between the two groups was detected.

Throughout the study, the median difference in positive-affect scores for each group when compared with baseline measurements was lower than 4 points, and at all time points, the 5th percentile and 95th percentile of difference contained the value zero which suggests that there is not enough evidence to show a significant difference. On average throughout the study, both test and control groups had decreased positive-affect scores compared with baseline levels; at the end of the study, positive-affect scores were 7 points and 6.5 points for the test and control groups, respectively (Figure 4(b); Supplementary Table 2).

Overall, the negative-affect scores at each time point were lower in the control group except at T6. At no point

within this study were significant differences in negative-affect score between participants using the ovulation test and the controls observed (Figure 5(a); Supplementary Table 1). The mean difference in negative-affect results was above zero for the majority of the time with a maximum difference of 3.33 points (95% CI: $-0.50, 7.16$; $P=0.0863$) being observed at T2; however, these results were not significant.

Throughout the study, the median difference between randomization groups in negative-affect scores compared with baseline measurements was lower than 3 points, and at all time points, the 5th percentile and 95th percentile of difference of negative-affect scores compared with baseline measurements contained the value zero for test and control groups. At the end of the study, both test and control group had a median increase of 1 point in negative affect compared with baseline levels (Figure 5(b); Supplementary Table 2).

SF-12v2. The SF-12v2 scores are split into physical and mental attributes, and no significant differences between randomization groups were found for either domain at baseline (mean difference; physical scale= -2.17 (95% CI: $-4.92, 0.58$; $P=0.12$); mental scale= -1.26 (95% CI: $-2.87, 5.39$; $P=0.48$)) or at the end of the study (mean difference; physical scale= 2.00 (95% CI: $-3.82, 7.83$; $P=0.54$); mental scale= -3.15 (95% CI: $-9.77, 3.46$; $P=0.33$)) (Supplementary Table 1).

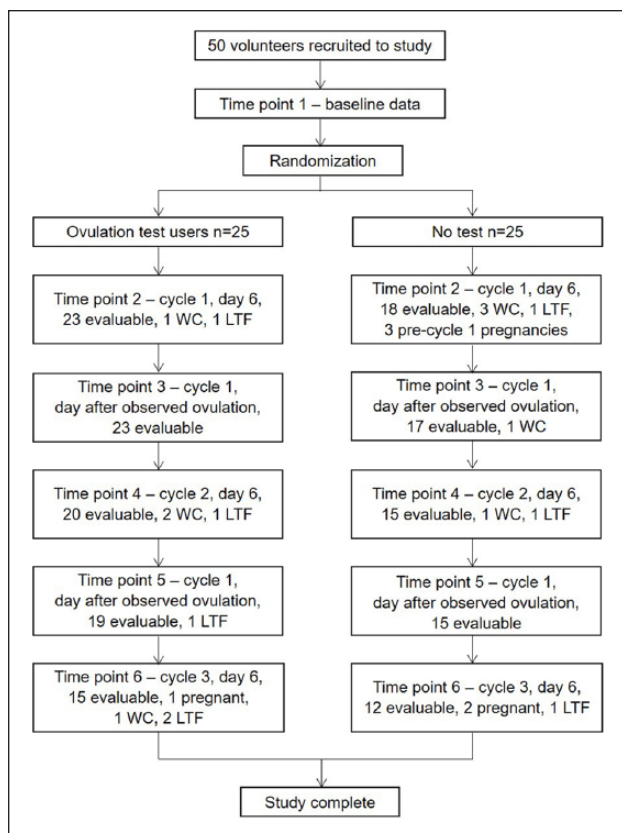


Figure 2. Schedule of events and numbers of volunteers at each time point during the study. The 50 volunteers were equally randomized into two groups; however, there was unequal loss of volunteers between groups for the duration of the study. “Pre-cycle 1 pregnancies” indicates that pregnancy occurred post-randomization but before their first study cycle was due to begin.

WC: withdrawn consent; LTF: lost to follow-up.

When scores were broken down into their individual components, no significant differences were observed within any components of the SF-12v2 between participants who used the ovulation test and those who did not.

EHIQ. The total score obtained from the EHIQ for the two randomization groups was compared using data collected at baseline and at the end of the study; no significant differences between the two groups were observed at either time point (Supplementary Table 1).

When the total EHIQ scores were broken down into their individual components, there was only one instance in which a significant difference between the test and control groups was observed. At T6, there was a significant difference within the category “couple concordance,” with ovulation test users having a concordance score of 25 and control group having a concordance score of 3.75. This amounts to a difference of 21.25 (95% CI: 9.25, 33.25; $P=0.0015$), which remained significant after adjusting for multiple testing within the EHIQ breakdown scoring using the Bonferroni method ($P=0.0268$).

Cortisol concentrations in the test and comparison groups. Whereas the log ratio of cortisol to creatinine was greater in the test group than in the control group at the beginning of the study, at T5 and T6, it was higher in the control group than in the test group. None of the differences for log cortisol to creatinine was statistically significant at any time point (maximum difference observed was -0.28 units at T5 (95% CI $-0.69, 0.13$; $P=0.1708$)). Likewise, for log E3G to creatinine, no significant difference was observed at any time point between participants using the ovulation test and those who did not. The maximum difference of 0.30 units was again observed at T3 (95% CI, $-0.15, 0.75$; $P=0.1850$). Supplementary Table 3 shows the means and standard deviations of the log ratio of cortisol to creatinine and E3G to creatinine over time for the test and the control groups (samples were collected at T2–T6). Graphical representations of these data are shown in Figure 6(a) and (b).

Correlation between total stress and other measures. Pearson’s correlation (r) between total stress and baseline demography was analyzed and showed some significant relationships, such as length of time trying to conceive ($r=0.467$; $P<0.0001$), previous live births ($r=0.336$; $P<0.0001$), and hours of exercise per week ($r=-0.299$; $P<0.0001$). Other baseline demographics measured, such as age ($r=-0.095$; $P=0.1759$), showed little evidence of a linear trend with the level of perceived stress.

There seemed to be no relationship between total stress (as assessed by PSS) and log ratio of E3G to creatinine levels within this study ($r=0.049$; $P=0.5609$). However, there was a relationship between log ratio of cortisol to creatinine and total stress score ($r=0.276$; $P=0.0008$).

Discussion

This study used multiple approaches to assess levels of stress in women with unexplained infertility, including a range of validated questionnaires and urinary biomarkers. No significant difference was observed in any of these measures between women who were using ovulation tests and women in the control group who were not using any such methods to identify their day of ovulation. This study therefore does not support the hypothesis that the use of digital ovulation tests may increase levels of clinically relevant stress in women with unexplained infertility who are about to begin fertility treatment.

All questionnaire results remained relatively constant over the duration of the study, and the results were comparable to published normative data in female populations. These data are also similar to previously published data from a cohort of women with no known evidence of infertility.¹³ Women with unexplained infertility were on average approximately 4 years older than women with no known fertility problems and with similar self-reported menstrual cycle lengths (28.7 and 30.5 days, respectively). The length of time trying to conceive was unsurprisingly longer in the

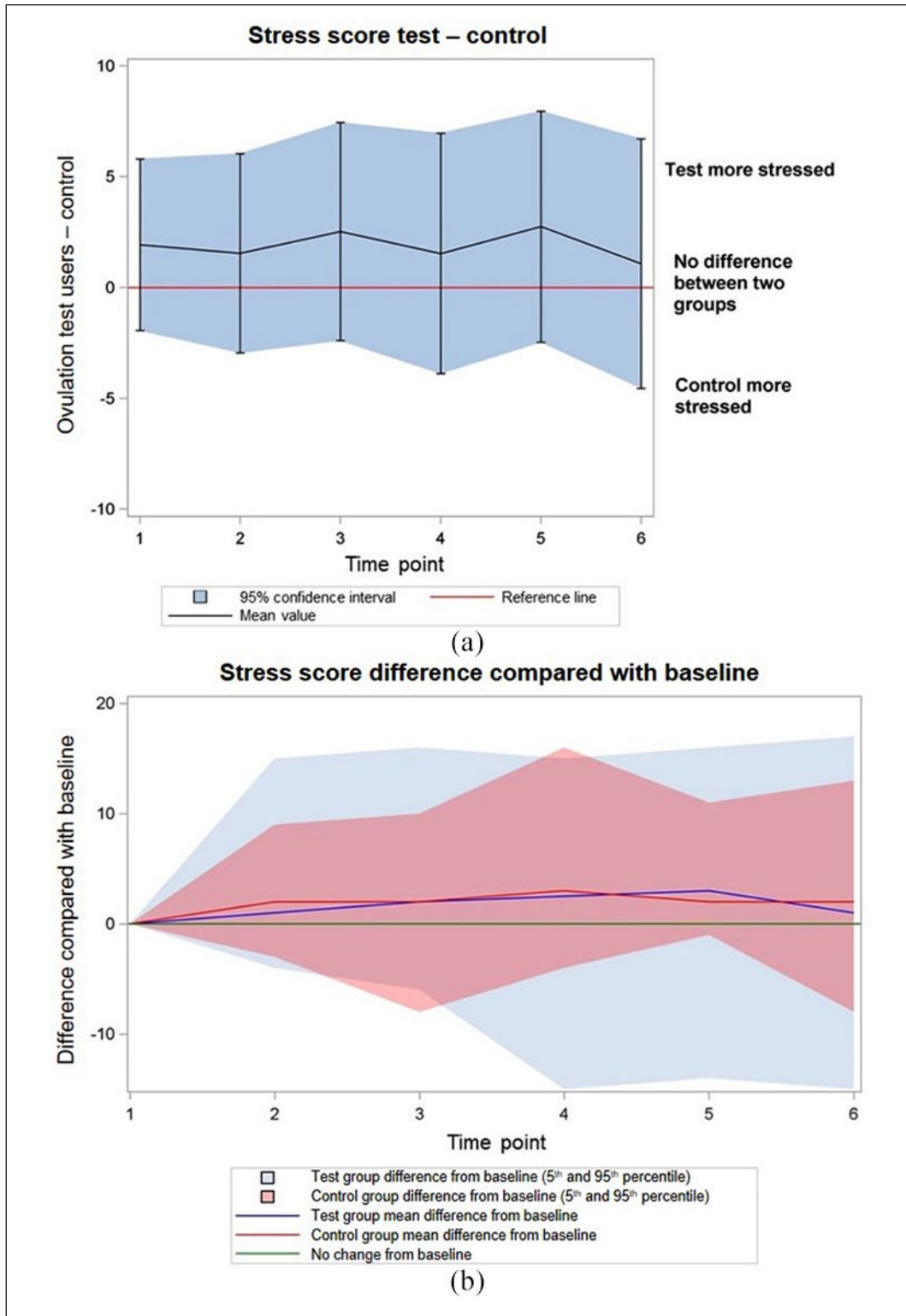


Figure 3. Graphical representations of total stress scores as measured using the Perceived Stress Scale questionnaire: (a) The mean difference in total stress scores comparing test group with control group and (b) the median, 5th and 95th percentile changes in stress score since baseline by randomization group.

unexplained-infertility population (22.8 months) than in the population with no known fertility problems (8.7 months).

With regard to other demographic data, the two populations were very similar.

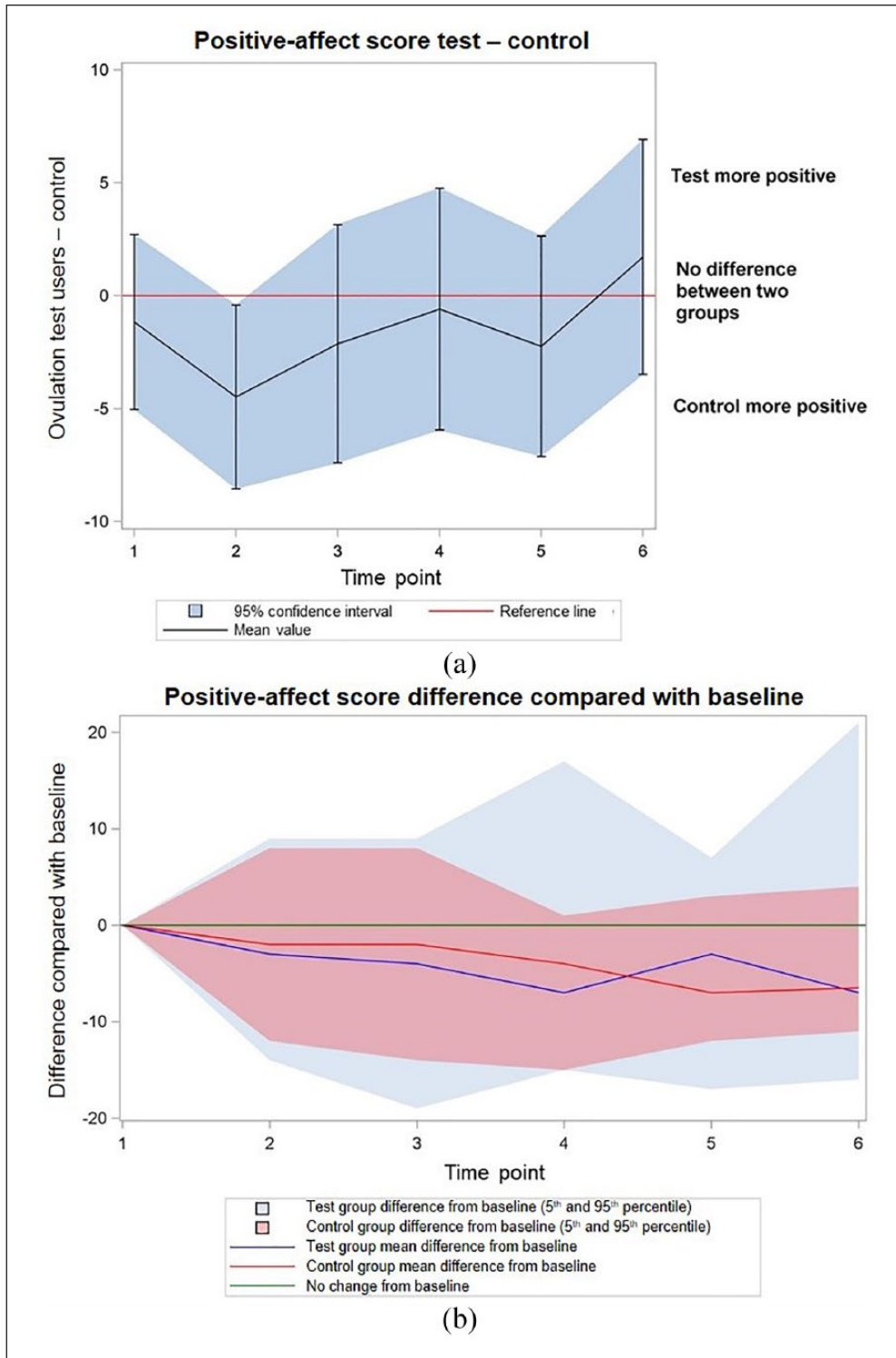


Figure 4. Graphical representations of positive-affect scores as measured by the positive domains of the PANAS: (a) The mean difference in positive-affect scores comparing test group with control group and (b) the median, 5th and 95th percentile changes in positive-affect score since baseline by randomization group. PANAS: Positive and Negative Affect Schedule.

Analysis of the correlations between levels of total stress and demographic data has shown that variables other

than randomization group may have a linear impact on the stress of a participant. In particular, an increased length in

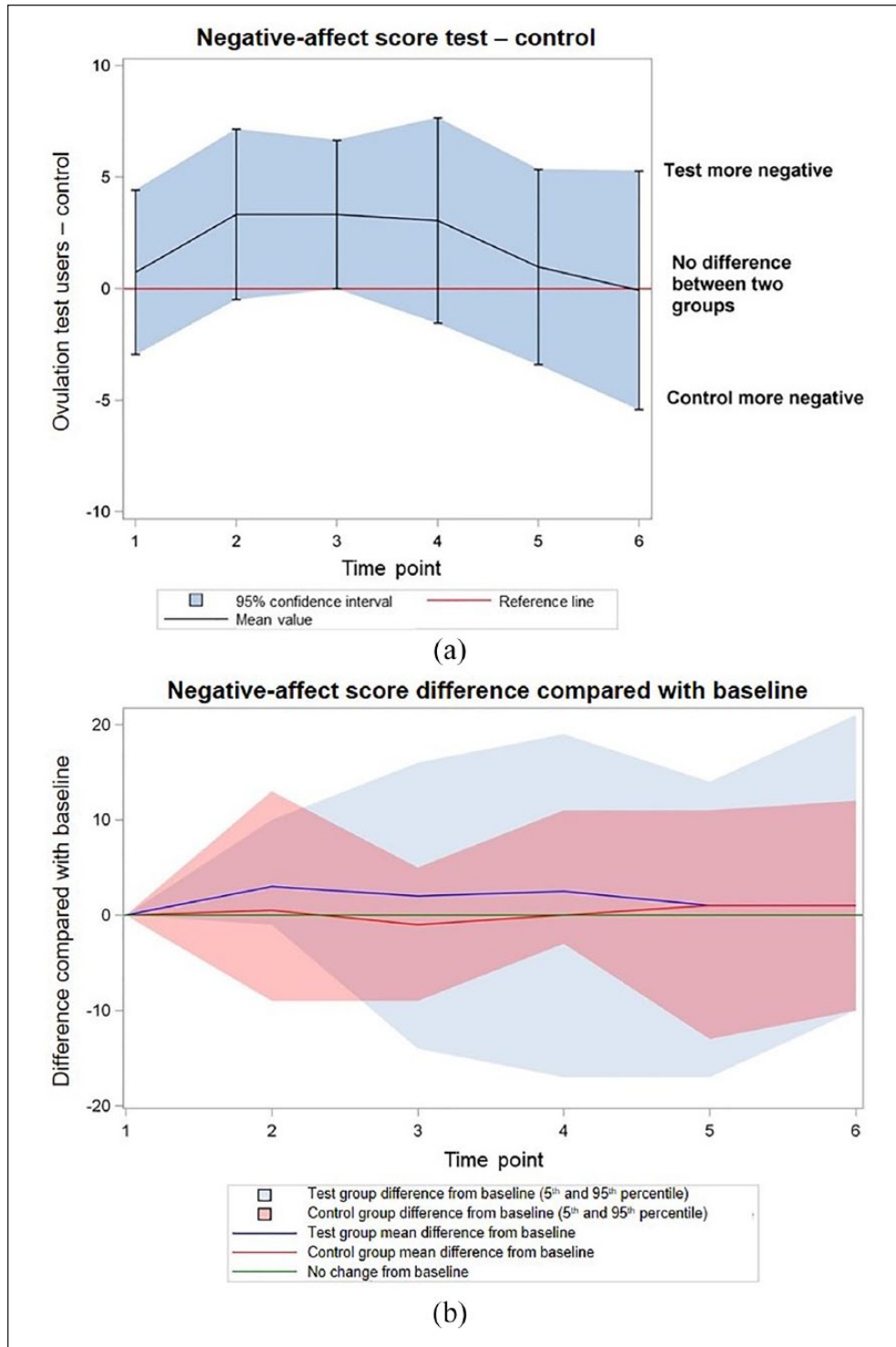


Figure 5. Graphical representations of negative-affect scores as measured by the negative domains of the PANAS: (a) The mean difference in negative-affect scores comparing test group and control group and (b) the median, 5th and 95th percentile changes in negative-affect score since baseline by randomization group. PANAS: Positive and Negative Affect Schedule.

time trying to conceive, a higher number of previous live births, and lower average hours of exercise per week correlated with an increase in the level of stress exhibited.

The pregnancy rate in this study was 12%, compared with 36.7% in the previous study population who had no known fertility problems. This suggests that the women

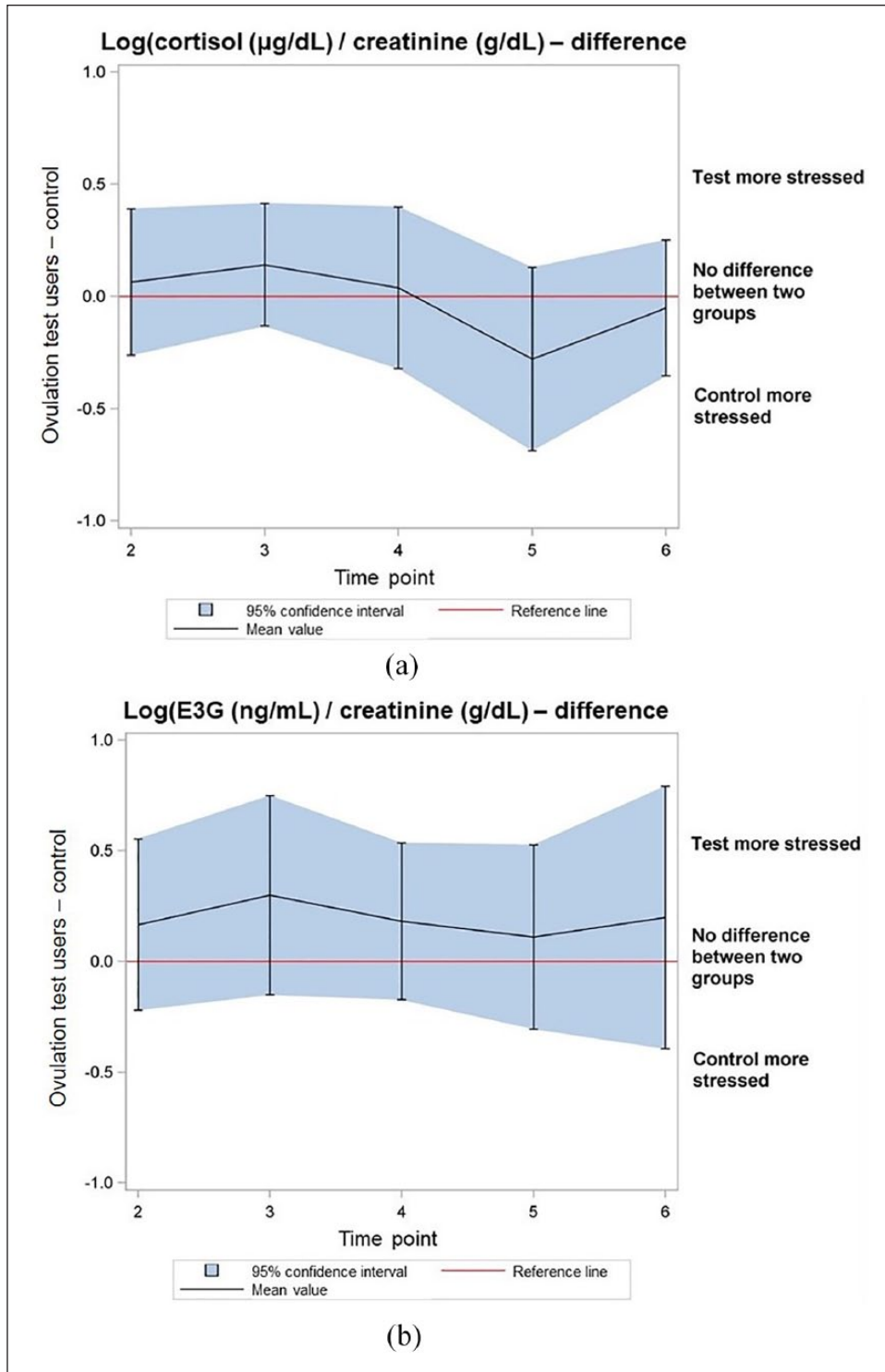


Figure 6. Graphical representations of biochemical marker data: (a) The mean difference in log cortisol: creatinine ratio scores comparing test group with control group and (b) the mean difference in log E3G: creatinine ratio scores comparing test group with control group.
E3G: estrone-3-glucuronide.

who were attending the fertility clinic had been correctly referred and were likely to require assistance to help them conceive, which is beyond what an ovulation test can offer.

In addition, unsuccessful prolonged use of ovulation tests should prompt further clinical investigation of patients as it could be a signal that there is an underlying fertility

problem. Indeed, this study did not examine whether prolonged use of ovulation tests could cause stress, so our data cannot be extrapolated to this scenario. Further studies would be needed to examine the impact of prolonged use on stress levels and emotional well-being.

Overall, use of digital ovulation tests by sub-fertile women under medical care had negligible negative effects and no detectable positive benefit on psychological well-being according to both questionnaire-based and biochemical measurements of stress. The significant result obtained on the sixth day of the first menstrual cycle on positive affect coupled with a slightly elevated log ratio of cortisol to creatinine at the beginning of the study suggests that there may be increased anxiety when first using the ovulation test, but these differences were insignificant at all subsequent time points.

The results obtained from breaking down the EHIQ data into the relevant domains suggest that for the patient population enrolled in the study, couple concordance was significantly lower in the control group than in the test group at the end of the study. Low scores in this particular domain of the questionnaire specifically suggest that couples disagree in terms of how much money to spend on their fertility treatment and their overall commitment to having children.

Couple concordance rose from a score of 22.83 at T1 to a score of 25 at T2 in the test group, which suggests a slight improvement in emotional health related to this domain. In comparison, couple concordance fell from a score of 10.94 at T1 to a score of 3.75 at T2 in the control group, showing a much lower emotional state in relation to this domain for those volunteers randomized to the control group. This suggests that couples in this study who were using the ovulation test were in closer agreement in terms of their fertility journey than those who were not using it. This may be a result of using the ovulation test itself, as it may have provided couples with the feeling that their decisions are based on personalized information giving them something to focus on together and a reason to engage with each other on the subject of treatment and trying to conceive. Indeed, use of ovulation tests was reported as having a positive impact on the relationship for some couples who took part in the main study,²⁰ and it would be interesting for future studies in this field to investigate this theme further.

The main limitation of this study is the small sample size, as a larger number of participants would have increased the power of the tests performed to look for smaller differences. However, the longitudinal nature of the study together with the multiple methods of measurement adds to the robustness of the findings. The study did show important relationships between demographics and stress, which demonstrate study numbers were sufficient to find important relationships, indicating that any effect of home ovulation test use (either positive or negative) was of less relevance to stress levels than other factors such as

length in time trying to conceive and number of previous live births. The length of study was appropriate for the patient population given the fertility clinic setting, and changes in clinically relevant levels of stress should be apparent over this period of time, although we cannot rule out that small differences in levels of stress could have existed between groups. In addition to the small sample size, there was also a relatively high number of patients who were either lost to follow-up or withdrawn from the study before completion (26.2%). As the numbers were equal between the two groups, we do not believe that this impacts the findings of the study but it does highlight the difficulty in recruiting and maintaining women who are trying to conceive in a medical setting. This emotionally difficult and sensitive experience can make it difficult for women to remain motivated to continue participation in clinical research and should be taken into consideration for future studies in this field.

Home ovulation tests have been found to have utility in both the clinical and home environment for women hoping to conceive.¹³ Users have reported feeling that by using ovulation testing, they are taking action and control of the process of trying to conceive and have an improved understanding of menstrual cycle phases.¹³ In previous publications, clinic-related waiting periods, such as those experienced when first being referred to fertility specialists, can be difficult and were linked with an increase in anxiety and depression.^{21,22} It is our belief that ovulation test usage at this time could be of benefit to patients, as it may give them a sense of self-control.

Conclusion

No clear evidence of an association between ovulation tests and levels of stress, psychological well-being, and quality of life in women with unexplained infertility was observed in this study. Therefore, this study does not support arguments that using digital ovulation tests can cause stress in women trying to conceive.

Acknowledgements

Professor Mike Campbell, University of Sheffield, provided assistance with the study design and statistical advice. Dr Debra Scates, Integrated Medhealth Communications (funded by SPD Swiss Precision Diagnostics Development Company Limited), provided medical writing assistance in the form of manuscript journal styling. Clinical trial registration number: NCT01084304; Assessment of Stress Levels in Consumers Using Ovulation Tests in Order to Conceive. <https://clinicaltrials.gov/ct2/show/NCT01084304>.

Author contributions

S.D. contributed to conducting the study. C.H. conducted the statistical analysis. W.L. contributed to the study design, interpretation of findings, and manuscript preparation. S.J. contributed

to the study design, interpretation of findings, and manuscript preparation. G.L.J. contributed to the study design, interpretation of findings, and manuscript preparation. S.W. contributed to the study design, and conduct, interpretation of findings, and manuscript preparation.

Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: S.W., C.H., and S.J. are employees of SPD Swiss Precision Diagnostics Development Company Limited, which provided funding for this study. S.D. has no conflict of interest to declare. G.L.J. has received a grant from SPD Swiss Precision Diagnostics Development Company Limited. W.L. is a Board Member of the Royal Hospital for Women Foundation and Flinders Fertility, he receives research support from Ferring Pharmaceuticals, Merck, and Swiss Precision Diagnostics, and is a minority share-holder in Virtus Health.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Funding for this study was provided by SPD Swiss Precision Diagnostics Development Company Limited. S.W., C.H., and S.J. are full-time employees of SPD Swiss Precision Diagnostics Development Company Limited.

Supplemental material

Supplemental material for this article is available online.

ORCID iD

Cameron Hogg  <https://orcid.org/0000-0002-7282-8160>

References

- Human Fertilisation & Embryology Authority. Fertility treatment 2014: trends and figures, <https://ifqlive.blob.core.windows.net/umbraco-website/1783/fertility-treatment-2014-trends-and-figures.pdf> (2014, accessed May 2018).
- Bernstein J, Brill M, Levin S, et al. Coping with infertility: a new nursing perspective. *NAACOGS Clin Issu Perinat Womens Health Nurs* 1992; 3(2): 335–342.
- Van Balen F and Trimbos-Kemper TC. Long-term infertile couples: a study of their well-being. *J Psychosom Obstet Gynaecol* 1993; 14(Suppl.): 53–60.
- Severy LJ, Robinson J, Findley-Klein C, et al. Acceptability of a home monitor used to aid in conception: psychosocial factors and couple dynamics. *Contraception* 2006; 73(1): 65–71.
- Alesi R. Infertility and its treatment—an emotional roller coaster. *Aust Fam Physician* 2005; 34(3): 135–138.
- Robinson JE and Ellis JE. Mistiming of intercourse as a primary cause of failure to conceive: results of a survey on use of a home-use fertility monitor. *Curr Med Res Opin* 2007; 23(2): 301–306.
- Mu Q and Fehring RJ. Efficacy of achieving pregnancy with fertility-focused intercourse. *MCN Am J Matern Child Nurs* 2014; 39(1): 35–40.
- Stanford J, White G and Hatasaka H. Timing intercourse to achieve pregnancy: current evidence. *Obstet Gynecol* 2002; 100(6): 1333–1341.
- Bouchard TP, Fehring RJ and Schneider MM. Achieving pregnancy using primary care interventions to identify the fertile window. *Front Med* 2017; 4: 250.
- Khattab AF, Mustafa FA and Taylor PJ. The use of urine LH detection kits to time intrauterine insemination with donor sperm. *Hum Reprod* 2005; 20(9): 2542–2545.
- Kopitzke EJ, Berg BJ, Wilson JF, et al. Physical and emotional stress associated with components of the infertility investigation: perspectives of professionals and patients. *Fertil Steril* 1991; 55(6): 1137–1143.
- NICE. *Fertility: assessment and treatment for people with fertility problems*. London: NICE, 2004.
- Tiplady S, Jones G, Campbell M, et al. Home ovulation tests and stress in women trying to conceive: a randomized controlled trial. *Hum Reprod* 2013; 28(1): 138–151.
- Cohen S, Kamarck T and Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* 1983; 24(4): 385–396.
- Watson D, Clark LA and Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol* 1988; 54(6): 1063–1070.
- Ware J, Kosinski M, Turner-Bowker D, et al. *How to score version 2 of the SF-12 health survey (with a supplement documenting version 1)*. Lincoln, RI: QualityMetric Incorporated, 2002.
- Nakamura K, Sheps S and Arck PC. Stress and reproductive failure: past notions, present insights and future directions. *J Assist Reprod Genet* 2008; 25(2–3): 47–62.
- Lokuge S, Frey BN, Foster JA, et al. Depression in women: windows of vulnerability and new insights into the link between estrogen and serotonin. *J Clin Psychiatry* 2011; 72(11): e1563–e1569.
- Matthews JN, Altman DG, Campbell MJ, et al. Analysis of serial measurements in medical research. *BMJ* 1990; 300(6719): 230–235.
- Jones G, Carlton J, Weddell S, et al. Women's experiences of ovulation testing: a qualitative analysis. *Reprod Health*. Epub ahead of print 21 December 2015. DOI: 10.1186/s12978-015-0103-y.
- Boivin J and Takefman JE. Stress level across stages of in vitro fertilization in subsequently pregnant and nonpregnant women. *Fertil Steril* 1995; 64(4): 802–810.
- Boivin J and Lancaster D. Medical waiting periods: imminence, emotions and coping. *Womens Health* 2010; 6(1): 59–69.