human reproduction

ORIGINAL ARTICLE *Infertility*

Home ovulation tests and stress in women trying to conceive: a randomized controlled trial

S. Tiplady[1],*, G. Jones[2], M. Campbell[2], S. Johnson[1], and W. Ledger[3]

[1]Scientific and Medical Affairs, SPD Development Company Limited, Bedford MK44 3UP, UK [2]School of Health and Related Research, University of Sheffield, Sheffield S14DA, UK [3]Obstetrics and Gynaecology, University of New South Wales, Sydney, NSW 2052, Australia

*Correspondence address. sarah.tiplady@spdspark.com

STUDY QUESTION: Does the use of a digital home ovulation test have any effect on the level of stress in women seeking to conceive?

SUMMARY ANSWER: No difference was found in levels of stress between women using digital ovulation tests to time intercourse compared with women who were trying to conceive without any additional aids: in addition, their use did not negatively impact time to conception in users but may provide additional benefits, including an increased understanding of the menstrual cycle, reassurance and confidence in focusing conception attempts to the correct time in the cycle.

WHAT IS KNOWN ALREADY: It has been suggested that timing of intercourse in such a way that it coincides with ovulation by using ovulation tests can lead to emotional distress; however, no study has been conducted to investigate this hypothesis specifically, until now.

STUDY DESIGN, SIZE AND DURATION: The study was performed over two complete menstrual cycles as a prospective, randomized, controlled trial including quantitative and qualitative methods. The intervention (test) group were given digital ovulation tests to time intercourse to the most fertile time of the cycle and the control group were provided with the current National Institute for Health and Clinical Excellence guidelines for increasing the chances of conception (intercourse every 2–3 days) and asked not to use any additional methods to time when ovulation occurs.

PARTICIPANTS/MATERIALS, SETTING AND METHODS: A total of 210 women who were seeking to conceive were recruited from the general UK population. A total of 115 women were randomized to the test group and 95 to the control group through block randomization. The positive and negative affect schedule (PANAS) and the Perceived Stress Scale (PSS) were used to measure subjective stress levels, the Short-Form 12 health survey was used as a measure of general health and well-being and urine samples were measured for biochemical markers of stress including urinary cortisol. Qualitative data were collected in the form of a telephone interview upon study completion.

MAIN RESULTS AND THE ROLE OF CHANCE: There was no evidence for a difference either in total stress as measured using the PSS or in total positive or negative affect using the PANAS questionnaire between the test and control groups at any time point for the duration of the study. During cycle 1, for example, on Day 6, the difference in total stress score (test−control) was −0.62 [95% confidence interval (CI) −2.47 to 1.24] and on the day of the LH surge, it was 0.53 (95% CI −1.38 to 2.44). In addition, no correlation was observed between time trying to conceive and levels of stress, or between age and levels of stress, and no evidence was found to show that stress affected whether or not a pregnancy was achieved. There is also no evidence that the biochemistry measurements are related to whether a pregnancy was achieved or of a difference in biochemistry between the treatment groups. The use of digital ovulation tests did not negatively affect time to conception and with an adequately sized study, could potentially show improvement. To ensure that the results of this study were not affected by chance, we used a number of different methods for measuring stress, each of which had been independently validated.

LIMITATIONS AND REASONS FOR CAUTION: Randomization occurred before the start of the study because of the need to provide the ovulation tests in readiness for Day 6 of the first cycle. As a consequence, a number of women fell pregnant during this period (22 and 13 in the test and control groups, respectively). A further 15 women were either lost to follow-up or withdrew consent prior to study start. Pregnancy rate was higher overall in the test group, so to ensure that there were sufficient data from women who failed to become pregnant in the test group, we implemented an additional biased recruitment. This second cohort may have been different from the first, although no significant differences were observed between the two phases of recruitment for any of the information collected upon admission to the study.

Introduction

Becoming pregnant is often not straightforward and it has been estimated that one in seven couples in the UK suffer from difficulty in conceiving (NICE, 2004). There are a myriad of causes for failing to conceive quickly, but a simple factor that is easily corrected is ensuring that intercourse occurs during the fertile window. Conception is most likely to occur when intercourse takes place on the day before or on the day of ovulation (Wilcox *et al.*, 2000) and studies have shown that with fertility-focused intercourse, conception is likely to occur more quickly (Hilgers *et al.*, 1992; Gnoth *et al.* 2002; Stanford, *et al.*, 2002; Robinson *et al.*, 2007). There are two main reasons why couples may incorrectly time intercourse, firstly because they are unaware of the most appropriate time in their cycle (Zinaman *et al.,* 2012) and secondly, the clinical prediction that the fertile window usually occurs between Days 10 and 17 of the menstrual cycle is often not correct (Wilcox *et al.*, 2000; Stanford *et al.*, 2002; Robinson and Ellis, 2007). This is because although normally cycling women have a mean cycle length of 27–28 days, the range both within and between individuals is considerably longer and has been reported to be as high as 36 days (Creinen *et al.*, 2004; Johnson *et al.*, 2009). Home ovulation tests are a simple and popular method for timing intercourse via the daily monitoring of urinary LH levels in order to detect the LH surge that occurs ∼24–36 h prior to ovulation (WHO, 1980; Singh *et al.*, 1984; Collins, 1985; Behre *et al.*, 2000).

It has been suggested that timing of intercourse so as to coincide it with ovulation by using ovulation tests or other fertility awareness-based methods can lead to emotional distress. The current National Institute for Health and Clinical Excellence (NICE) guidelines state that couples who are concerned about their fertility should be informed that sexual intercourse every 2–3 days optimizes the chances of pregnancy. The guidelines also state that timing of intercourse so as to coincide it with ovulation causes stress and is not recommended (NICE, 2004). The evidence supporting this statement in the guidelines is limited to a retrospective observational study of 26 patients undergoing infertility treatment, where it was found that events associated with lack of conception or loss of pregnancy, such as a negative pregnancy test, onset of menses, ectopic pregnancy or miscarriage, were in fact emotionally more difficult (Kopitzke *et al.*, 1991). Indeed, failure to become pregnant is likely to be the greatest cause of stress when a woman is trying to conceive (Severy *et al.*, 2006). It is also reasonable to suggest that attempts to undertake frequent intercourse for the duration of the cycle (NICE, 2004) may prove excessively demanding and thus as emotionally challenging as timing intercourse to the fertile period.

The advancement in ease of use and accuracy of home ovulation test devices, in particular the introduction of digital displays, means that women now find it easier to interpret the results of these tests, which can in turn lead to greater certainty (Johnson *et al.*, 2011). In addition, many women have a poor understanding of their menstrual cycle and the timing of ovulation (Zinaman *et al.*, 2012). Home ovulation test devices can help build awareness and knowledge of oneself during attempts to conceive as well as help to empower and engage women to take control of their own fertility (Brown *et al.*, 1987; Blackwell *et al.*, 2003).

We conducted this study in response to the misunderstanding and confusion surrounding whether such products promote stress. The aim of this study was to compare levels of stress in women seeking to conceive while using digital home ovulation tests compared with women who were provided with the NICE clinical guidelines on increasing the chances of conception. We hypothesized that there would be no measurable difference in the level of stress associated with either method of trying to conceive as for both intercourse is determined by instruction and not necessarily by choice.

Stress and negative emotions can be expressed via biochemical and/or behavioural channels, which allow them to be measured in a number of ways including psychological questionnaires, biochemical markers and qualitative interview techniques. Questionnaires are useful tools to measure subjective phenomena such as symptoms and quality of life and biomarkers are chemical indicators of biological state. Both are useful tools for measuring baseline information and in evaluating change over time while qualitative techniques are useful in exploring common themes arising from in-depth discussions. Cortisol is a well-established biomarker associated with increased levels of stress (Miki and Sudo, 1998; Nicolson, 2008; Nepomnaschy *et al.*, 2011), and there are several psychological questionnaires that are well-validated and commonly used for observing both specific and non-specific stress in clinical settings. All three of these measures were incorporated into this randomized, controlled study.

Materials and Methods

Study design and recruitment

The study was performed as a prospective, randomized controlled trial including quantitative and qualitative methods. The study protocol was approved by SPD ethics committee. Volunteers were recruited via an advert placed on the Clearblue UK website, which attracted 550 responses. Detailed information about the study was provided to the

volunteers and written informed consent was obtained before the commencement of the study.

A total of 210 volunteers were recruited from the UK population (21 February–5 December 2010). The inclusion criteria were women living in the UK who were aged between 18 and 40 years, having regular menstrual bleeds and wishing to become pregnant. Excluded were women who had been using hormonal contraception in the last 3 months, women currently undergoing fertility treatment or investigation, women who had previously been diagnosed as infertile, anyone with a history of depression, anxiety or panic attacks and anyone dependant on either drugs or alcohol. Women who had previously used ovulation tests were not excluded from participating in the study.

Study population, sample size and randomization

Since the variability of stress was unknown, a pragmatic decision was taken to recruit 75 subjects per group, which was feasible in the time period and would give 80% power to detect a difference of 0.4 SDs at 5% significance level, which is a moderate effect size. Kopitzke *et al.* (1991) demonstrated a difference in stress in women trying to conceive based on a sample size of 26. Initially, 150 volunteers were recruited to the study and these were randomized equally either into the test or into the control group through block randomization. Test-group volunteers used the Clearblue Digital Home Ovulation Test for the duration of the study. They were asked to begin testing on Day 6 of their cycle regardless of their normal cycle length. Control-group volunteers were asked not to identify their time of ovulation using methods such as ovulation testing or basal body temperature measurements and instead were advised of the NICE guidelines on how to increase the chances of conception, i.e. that sexual intercourse every 2–3 days for the duration of the cycle is likely to increase the chances of conception. It was the volunteers' choice as to whether or not they followed these guidelines. On completion of the study, the control-group volunteers were provided with Clearblue Digital Home Ovulation Tests as an incentive for complete participation.

Upon completion of the study by the initial 150 volunteers, it was found that there were more pregnancies than expected (44 and 26 in total for the test and the control groups, respectively). This resulted in fewer questionnaires and urine samples being available from ovulation test users who failed to become pregnant while on the study, so that there was insufficient power to ascertain a true representation of levels of stress in this group compared with the control group. Therefore, an additional cohort (60 volunteers in total) was recruited in the same way as the main study, only randomization was weighted at a ratio of 2:1 into the test group in order to enrich the data in this group. This resulted in a final study population of 210 volunteers with 115 randomized to the test group and 95 to the control group.

Randomization schedules were generated using the STATA software by a statistician who was not directly involved in the recruitment of the study. Un-blinding of the study did not take place until statistical analysis of the data was complete.

Study protocol

Upon admission to the study, a brief medical history was collected from each volunteer, including demographic information, obstetric and gynaecological history as well as information regarding their general health. Once recruited, volunteers were assigned to a randomization group by the study co-ordinator, sent the required study materials and asked to begin in their next menstrual cycle. For some volunteers, this resulted in a pregnancy being achieved prior to the start of Cycle 1. These were known as 'Pre-cycle 1 pregnancies'. Volunteers were asked to collect first morning urine samples and to complete a range of questionnaires at specific time points over two complete menstrual cycles according to the schedule presented in Fig. 1. Levels of stress were determined both biochemically and by questionnaire. The additional measurement at time point 6 in Cycle 3 was included to ensure that all participants completed two full cycles, while on the study and to ensure that pregnancy rate in Cycle 2 could be captured. This time point also allowed levels of stress to return to baseline where necessary after receiving the news that pregnancy had not been achieved after two cycles of trying to conceive while on the study.

Quantitative data collection

The questionnaires used were the Perceived Stress Scale (PSS) (Cohen *et al*, 1983), the positive and negative affect schedule (PANAS) (Watson *et al.*, 1988) and the Short Form-12 Health Survey (SF-12) (Ware and Sherbourne, 1992). The biochemical marker of stress was urinary cortisol, which is a well-established stress indicator (Nakamura *et al.*, 2008). In addition, urinary estrone-3-glucuronide (E3G) was analysed as a marker of estrogen status as levels of this marker are known to be associated with depression and anxiety (Lokuge *et al.*, 2011).

The PSS questionnaire is a measure of the degree to which situations in one's life are appraised as stressful. Items were designed to highlight how unpredictable, uncontrollable and overloaded respondents find their lives. The instrument uses a 5-point Likert scale rating from '0 = never' to '4 = very often'. Scores can range from 0 to 40, with higher scores indicating a higher level of perceived stress (Cohen *et al.*, 1983). The mean score for a normative sample of females completing the PSS questionnaire is 13.7 ± 6.6 (Cohen and Williamson, 1998).

The PANAS questionnaire consists of 10 positive affects (PA), (interested, excited, strong, enthusiastic, proud, alert, inspired, determined, attentive and active) and 10 negative affects (NA), (distressed, upset, guilty, scared, hostile, irritable, ashamed, nervous, jittery and afraid). Participants are asked to rate items on a scale from 1 to 5, on the basis of the strength of emotion where '1 = very slightly or not at all' and '5 = extremely'. For both of the PANAS domains, scores can range between 10 and 50. A higher score on the positive domain indicates greater PA (e.g. they are happier). However, a higher score on the negative domain indicates greater NA (e.g. they are more depressed) (Watson *et al.*, 1988). The validity and reliability of the PANAS questionnaire and its relationship between other measures of depression and anxiety was determined by Crawford and Henry (2004) where the mean score for a normative sample of females completing the PANAS questionnaire was 30.62 (SD 7.89) for PA and 16.68 (SD 6.37) for NA. The PSS and PANAS questionnaires were completed at time points 1–6 as shown in Fig. 1.

Finally, the SF-12 is a brief, well-validated and reliable generic questionnaire for capturing general health status and outcome information. The SF-12 health survey is one of the most widely cited measures of functional health status. The UK female population have been found to have scores of 49.54 and 49.17 for the physical and mental attributes, respectively, using the SF-36 version 2 questionnaire (Jenkinson *et al.*, 1999). The SF-12 questionnaire used in this study has been shown to yield comparable scores with the SF-36 questionnaire in both general and specific populations. It was completed at baseline (time point 1) by all study volunteers and again at the end of the study (time point 6) only by the volunteers who did not get pregnant.

First morning urine samples were collected at time points 2–6 using sample collection pots containing sodium azide as a preservative. Upon receipt of samples at the study site, they were added to a central sample database, aliquoted and stored at $-80°C$ until analysis. All the urine samples were analysed for cortisol and E3G using fully validated immunoassay systems and normalized for urine volume using creatinine measurements.
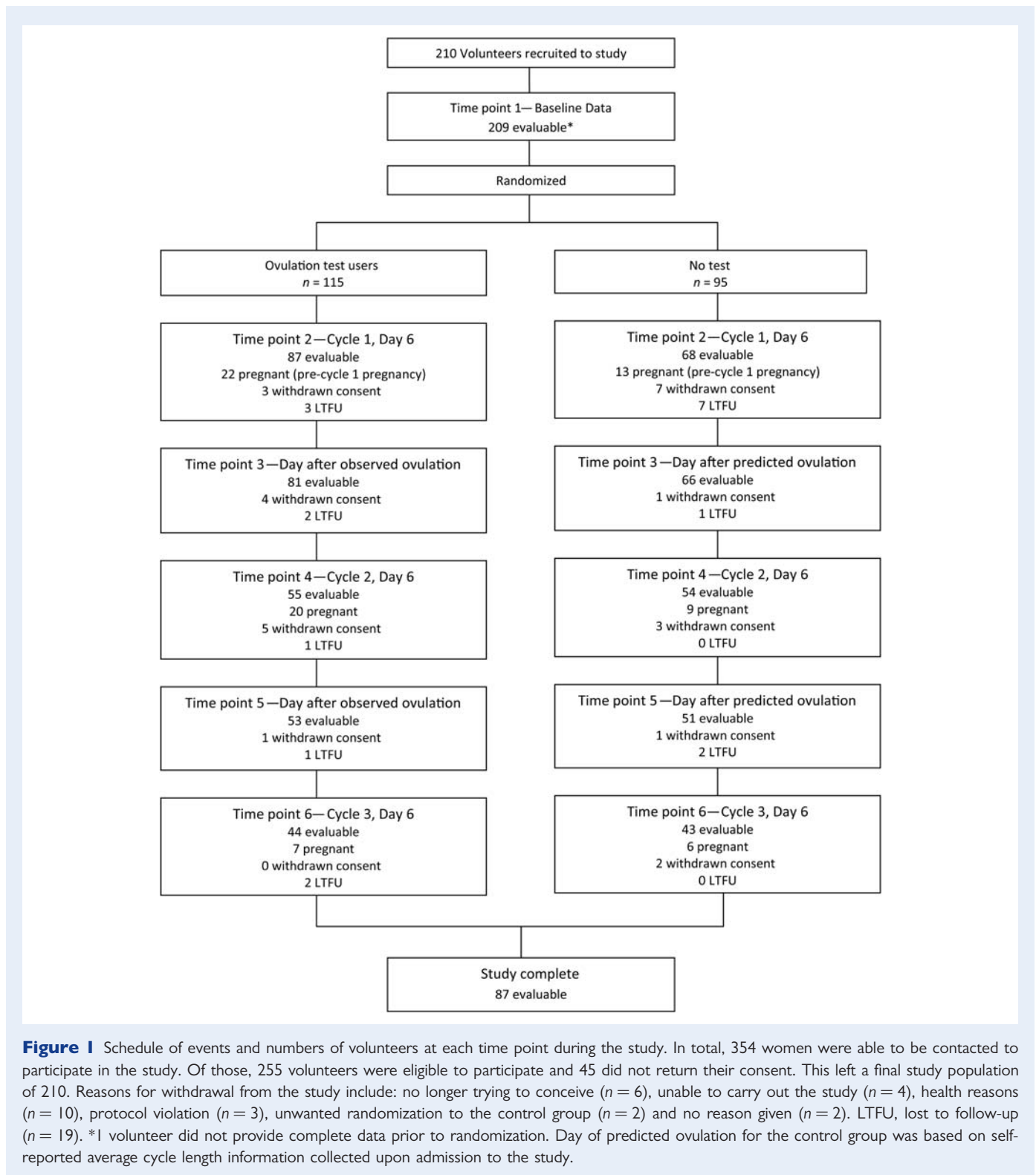
Figure 1 Schedule of events and numbers of volunteers at each time point during the study. In total, 354 women were able to be contacted to participate in the study. Of those, 255 volunteers were eligible to participate and 45 did not return their consent. This left a final study population of 210. Reasons for withdrawal from the study include: no longer trying to conceive ($n = 6$), unable to carry out the study ($n = 4$), health reasons ($n = 10$), protocol violation ($n = 3$), unwanted randomization to the control group ($n = 2$) and no reason given ($n = 2$). LTFU, lost to follow-up ($n = 19$). *1 volunteer did not provide complete data prior to randomization. Day of predicted ovulation for the control group was based on self-reported average cycle length information collected upon admission to the study.

Cortisol was measured using the Access 2 auto-analyser system from Beckman Coulter (High Wycombe, UK). The sensitivity of this urine assay is 0.4 µg/dl (11 nmol/l). Intra-assay % coefficient of variance (CV)s were 6.7, 4.4 and 4.4 and inter-assay %CVs were 7.9, 6.0 and 6.4 at low, medium and high concentrations, respectively. E3G analysis was carried out using a competitive immunoassay that was developed for urine samples by SPD for use on the DELFIA auto-analyzer system from Perkin Elmer (Waltham, Massachusetts) (Miro et al., 2005). The sensitivity of this urine assay is 0.076 ng/ml. Intra-assay %CVs were 2.1, 2.0 and 2.5 and inter-assay %CVs were 0.7, 0.9 and 1.0 at low, medium and high concentrations, respectively. Creatinine analysis was also carried out to enable volume correction of the markers indicative of stress. This measurement was made using the ABX Pentra 400 analyser from Horiba (Kyoto, Japan). The sensitivity of this urine assay is 1.39 mg/dl (123 µmol/l).

Intra-assay %CVs were 3.3, 0.6 and 0.5 and inter-assay %CVs were 6.0, 1.9 and 1.8 at low, medium and high concentrations, respectively.

Qualitative data collection

Upon completion of the study, individual, semi-structured telephone interviews were carried out with all the study participants to gain feedback on the use of the ovulation test, attitude to trying to conceive and further insight into their emotional wellbeing. The interview schedule comprised of both closed and open-ended questions. Interviews ranged from 10 to 30 min in length.

Data analysis of the interviews was carried out using a thematic approach (Marshall and Rossman, 1999), which is widely used in qualitative research. This enables data sources to be analysed in terms of the principal concepts or themes. These themes were then developed by the analyst, to enable the data to be reduced to key ideas.

Statistical analysis

The standardized psychological questionnaires (PSS, PANAS and SF-12) 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 baseline and remaining time point assessments (Matthews *et al.*, 1990). A mixed-effects model was applied to the scores over time. Sensitivity analyses were also performed to check the robustness of the conclusions, given the nature and number of missing data values in this study. Imputation methods were used to assess the effect of missing data due to attrition. Logistic regression was used to assess whether biochemical and psychological measures of stress can predict whether a subsequent conception will occur. Log transformation of the biochemical marker measurements was conducted to stabilize variance. The demographic data captured at recruitment were also analysed to examine the influence of age and time spent trying to conceive on stress levels over the duration of the study.

A principal components analysis (PCA) was also applied to the stress scores and to the biochemical measures of stress, categorizing the subjects on their demographic characteristics. Finally, Kaplan–Meier analysis was used to estimate time-to-event (i.e. conception) distributions in the two treatment groups and odds ratios were calculated.

Results

Study population

The mean age of 210 volunteers participating in the study was 28.87 years (range:19–40). The mean self-reported menstrual cycle length was 30.54 days (range: 25–49) and on average the study volunteers had been trying to conceive for 8.7 months (range: 0–84). Table I shows the various demographic variables by randomization group.

All the 210 women were randomized and of these 35 (16.7%) did not begin the study because of becoming pregnant prior to the start of Cycle 1 'Pre-cycle 1 pregnancy' and a further 20 withdrew from the study or were lost to follow-up before the study started as detailed in Fig. 1. Of the remaining 155 volunteers, 87 (56.1%) completed the study without achieving a pregnancy and 42 (27.1%) achieved a pregnancy either during Cycle 1 ($n = 29$) or during Cycle 2 ($n = 13$). The remaining 26 (16.8%) volunteers were lost to follow-up or withdrew from the study for various reasons as detailed in Fig. 1.

During recruitment, over a quarter of volunteers (28.1%) believed their cycle length to be the text book 28 days. Data available from all the volunteers up until the point at which they left the study were used.

Pregnancy rates

Table II gives the outcomes of the study by pregnancy status or withdrawal. For those randomized, 43% achieved a pregnancy in the test group, compared with 30% in the control. Excluding those who became pregnant and who were lost to follow-up before the start of the study, the rates were 31% and 22% for the test and the control groups, respectively. Kaplan–Meier estimates of time-to-event (i.e. conception) are shown in Figs 2 and 3 for all the subjects (intention-to-treat analysis) and for the subjects present at the start of the study (per protocol analysis) respectively.

Odds ratio was calculated for women who completed the study, excluding Pre-cycle 1 pregnancy individuals and those who were lost to follow-up prior to the start of the study. This analysis showed that the odds of getting pregnant in the test group was 1.59 [95% confidence interval (CI): 0.7652, 3.038] compared with the control group. Analysing the full study population (intention-to-treat), the odds of getting pregnant in the test group were found to be 1.77 (95% CI: 0.9992, 3.1585) compared with the control group.

Levels of stress as measured by questionnaire

In order to analyse the questionnaire data, comparisons were made between the women in the test group and those in the control group. Volunteers were categorized according to outcome: those who completed the study without getting pregnant, those who got pregnant in Cycle 1 of the study, those who got pregnant in Cycle 2 of the study, those who got pregnant before the study began and those who were exited from the study for other reasons, such as withdrawn consent or lost to follow-up (this outcome was labelled 'other'). Logistic regression analysis was used to assess whether stress can predict pregnancy. Results were analysed for total stress, which was determined using the PSS and for total PA and NA as determined by the PANAS questionnaire.

Table III shows summary statistics for total stress and total PA and NA scores at all time points of the study. There was no evidence of a difference in total stress, total PA or total NA between the test and the control groups or between the different outcome groups and no evidence of a change in levels over time. This finding was also supported by the PCA, which found no substantive relationships between randomization group, study outcome, time point or any indices of stress measured.

Figure 4 shows PCA analysis of the first two components fitted to all the data, including all the questions from the questionnaires, the demographic characteristics and the biochemical markers, identified by measurement type and labelled by time point. The first component identified a trend that is explained by the 'positive' and 'negative' questions/responses to the PANAS and PSS questionnaires and the second component identified a difference (or separation) between the SF-12 questions/responses and the other measurements. The 'positive' questions were negatively correlated with the 'negative' questions and appear on opposite sides of the graph. The questions from the SF-12 scale were grouped in the bottom half of the graph;

Table I Demographic variables by randomization group (mean, median and range except where indicated).

	Test group (n = 115)			Control group (n = 95)		
	Mean	**Median (SD)**	**Range**	**Mean**	**Median (SD)**	**Range**
Total previous pregnancies	1.7	1	0–7	1.8	2	0–7
Total live births	0.95	1	0–4	1.11	1	0–7
Total miscarriages	0.79	0	0–5	0.57	0	0–4
Months trying to conceive	8.6	6	0–60	8.8	6	1–84
Cycle length (days)	30.56	29	25–49	30.57	30	25–46
Height (m)	1.65	(0.071)	1.42–1.80	1.65	(0.066)	1.46–1.83
Weight (kg)	73.7	(15.82)	49.9–114.3	73.1	(18.92)	47.6–173.0
BMI (kg/m^2)	26.9	25.6	17.47–43.82	26.8	24.6	17.58–59.86
Alcohol (units/week)	2.41	0.50	0–15	2.11	0	0–10
Exercise (h/week)	3.0	2	0–20	3.5	3	0–21
Age (years)	28.3	28	20–40	29.7	30	19–39
Pre-cycle 1 Period (days)	30.76	23	9–243	29.75	27	5–127
Smoking History						
Yes		10 (8.7%)			4 (4.2%)	
Ex		30 (26.1%)			45 (21.4%)	
No		75 (65.2%)			76 (80.0%)	
Previous ovulation test use						
Yes		59 (51.3%)			60 (63.2%)	
No		56 (48.7%)			35 (36.8%)	

Table II Breakdown of volunteer outcome by pregnancy or withdrawal.

Frequency (column % age)	Test group (%)	Control group (%)	Total
Withdrew from Study	13 (11.30)	14 (14.74)	27
Lost to follow-up	9 (7.83)	10 (10.53)	19
Not pregnant	44 (38.26)	43 (45.26)	87
Pre-cycle 1 Pregnancy	22 (19.13)	13 (13.68)	35
Pregnant Cycle 1	20 (17.39)	9 (9.47)	29
Pregnant Cycle 2	7 (6.09)	6 (6.32)	13
Total	115	95	210

Pre-cycle 1 pregnancy—pregnancy achieved during cycle of consent, i.e. between consenting to the study and before starting Cycle 1.

the 'positive' questions were grouped together in the bottom left quadrant of the graph and the 'negative' questions were grouped together around the middle and to the left of the graph.

SF-12 scores are split into physical and mental attributes and summary statistics are also presented in Table III. No difference was found for either mental or physical attributes at time point 6 between the two randomization groups. In fact, the figures were similar for the test and control groups at the end of the study in the mean mental component (46.40 and 46.15, respectively) and the mean physical component (41.86 and 41.12, respectively).

The correlation between total stress and length of the time trying to conceive was also analysed at the start of the study as it was possible that a longer time in trying to conceive could lead to a higher level of stress. Results showed that for all time points the correlation was positive, but very low. The PCA did not support a correlation between stress and time trying to conceive. When analysing other demographic data collected upon admission, no correlation was observed between total stress and age, number of previous pregnancies or miscarriages, and no evidence was found to show that stress affects whether or not a pregnancy will be achieved.

Levels of stress as measured by biochemical marker

Figure 5 is the PCA analysis of the first two components fitted to all the data, except the individual questions, focusing on the biochemical markers, identified by measurement type and labelled by time point. There was no clear separation between the two randomization groups or between outcomes in either of the first two components fitted. The first component related to the stress score identifying a trend that is explained by the 'positive' and 'negative' questions/responses. The second component related to the correlations between the biochemical markers.

There was no evidence of a relationship between the demographic characteristics and outcome or treatment group, with the exception of alcohol consumption. High alcohol consumption was associated with the biochemistry measurements but only in a very few volunteers.

The means and standard deviations of the ratio of cortisol to creatinine and E3G to creatinine, over time, for the test and the control groups (samples were collected at time points 2–6) can be
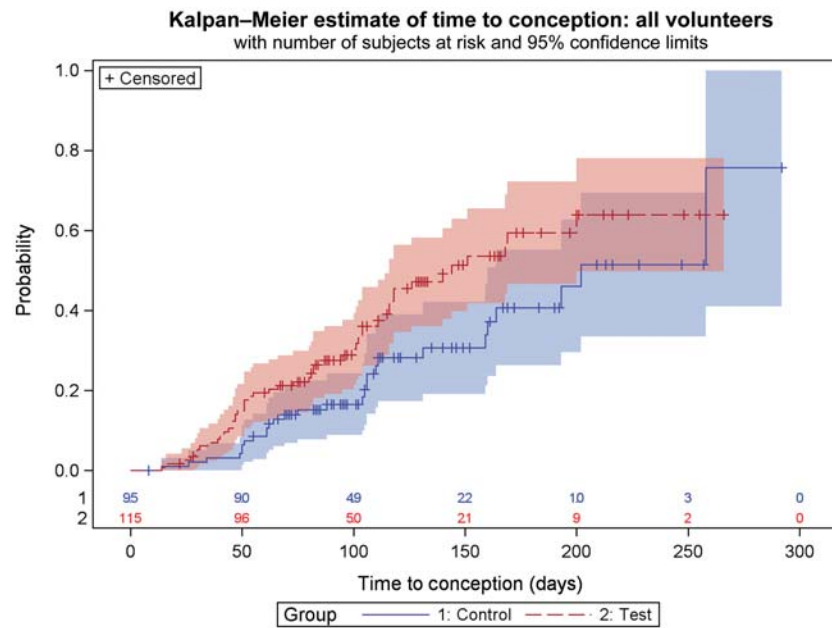
Figure 2 Kaplan–Meier estimate of time to conception for all volunteers recruited to the study with 95% confidence limits represented by the pink and blue areas surrounding the lines. Time to conception is based on time from recruitment to the study. Censored data indicate women who are lost to follow-up, for any reason, at the time indicated. Numbers at risk in each randomization group are given along the *x*-axis. Attrition of volunteers is due either to achieving a pregnancy or to completion of study.

Figure 3 Kaplan–Meier estimate of time to conception for volunteers excluding those that exited the study prior to Cycle 1. The 95% confidence limits are represented by the pink and blue areas surrounding the lines. Time to conception is based on time from recruitment to the study. Censored data indicate women who are lost to follow-up, for any reason, at the time indicated. Numbers at risk in each randomization group are given along the *x*-axis. Attrition of volunteers is due either to achieving a pregnancy or to completion of study.
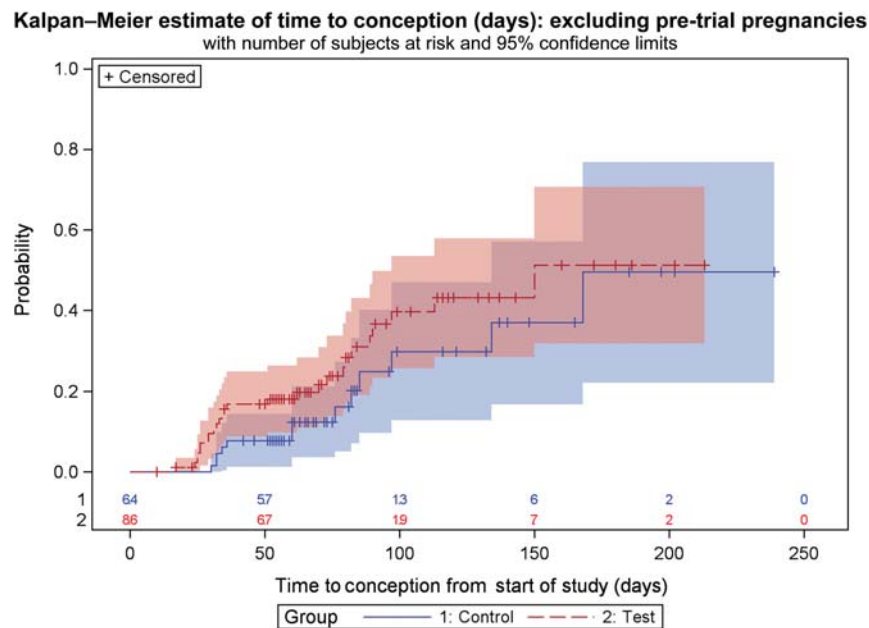
Table III Summary statistics (means and SDs) for all measurements at all time points; total stress (PSS), total positive and negative affect (PANAS), physical and mental attributes (SF-12) and cortisol: creatinine and E3G:creatinine ratios.

Total stress	Test			Control			Difference test-control	95% CI
	n	**Mean**	**SD**	**n**	**Mean**	**SD**		
T1—Baseline	114	14.99	4.58	93	14.80	5.10	0.19	−1.15 to 1.54
T2—Cycle 1, Day 6	81	16.11	4.92	63	16.73	6.04	−0.62	−2.47 to 1.24
T3—Cycle 1, Surge Day	73	16.79	4.86	61	16.26	6.08	0.53	−1.38 to 2.44
T4—Cycle 2, Day 6	52	17.29	6.00	51	16.10	5.49	1.19	−1.06 to 3.44
T5—Cycle 2, Surge Day	48	17.25	5.43	47	15.23	5.38	2.02	−0.19 to 4.22
T6—Cycle 3	37	17.76	6.48	40	15.78	6.25	1.98	−0.91 to 4.87
Negative affect								
T1—Baseline	113	14.16	4.03	93	14.37	4.58	−0.21	−1.41 to 0.99
T2—Cycle 1, Day 6	82	16.29	5.76	62	16.84	6.64	−0.55	1.06 to −2.64
T3—Cycle 1, Surge Day	73	15.85	5.02	62	15.39	5.02	0.46	0.87 to −1.25
T4—Cycle 2, Day 6	51	18.37	6.17	50	17.44	6.16	0.93	1.23 to −1.50
T5—Cycle 2, Surge Day	47	16.62	6.21	47	16.28	6.60	0.34	1.32 to −2.28
T6—Cycle 3	38	17.55	6.97	40	16.90	6.64	0.65	1.54 to −2.42
Positive affect								
T1—Baseline	114	36.14	7.00	92	37.96	6.26	−1.82	−3.64 to 0.01
T2—Cycle 1, Day 6	82	32.52	7.06	62	32.16	8.72	0.36	−2.32 to 3.05
T3—Cycle 1, Surge Day	72	32.58	8.06	62	34.63	7.55	−2.05	−4.72 to 0.62
T4—Cycle 2, Day 6	50	31.16	8.71	50	33.04	8.16	−1.88	−5.22 to 1.47
T5—Cycle 2, Surge Day	47	31.55	8.50	47	34.28	7.78	−2.73	−6.06 to 0.61
T6—Cycle 3	36	29.75	10.24	38	34.26	8.06	−4.51	−8.80 to −0.22
SF-12								
Physical T1—Baseline	115	39.80	3.53	93	39.65	3.28	0.15	−0.78 to 1.10
Physical T6—Cycle 3, Day 6	38	41.86	4.00	40	41.12	3.14	0.74	−0.87 to 2.37
Mental T1—Baseline	115	49.57	3.89	93	48.42	4.50	1.15	0.01 to 2.31
Mental T6—Cycle 3, Day 6	38	46.40	7.15	40	46.15	5.11	0.25	−2.54 to 3.04
Cortisol (μg/dl):Creatinine (g/dl)								
T2—Cycle 1, Day 6	84	136.29	62.82	66	157.23	138.92	−20.94	−57.35 to 15.46
T3—Cycle 1, Surge Day	69	129.15	70.36	61	149.04	85.20	−19.89	−47.22 to 7.44
T4—Cycle 2, Day 6	51	147.68	110.30	44	147.57	89.59	0.11	−40.63 to 40.85
T5—Cycle 2, Surge Day	47	127.75	65.42	45	149.16	92.40	−21.41	−54.70 to 11.88
T6—Cycle 3	37	139.33	59.03	38	156.23	89.44	−16.9	−51.69 to 17.89
E3G (ng/ml):Creatinine (g/dl)								
T2—Cycle 1, Day 6	84	98.99	56.69	66	101.34	46.23	−2.35	−18.94 to 14.26
T3—Cycle 1, Surge Day	69	403.15	237.43	61	276.31	180.57	126.84	54.10 to 199.58
T4—Cycle 2, Day 6	51	104.90	79.52	44	108.32	56.13	−3.42	−31.19 to 24.35
T5—Cycle 2, Surge Day	47	311.11	206.71	45	261.54	182.65	49.57	−31.14 to 130.28
T6—Cycle 3	37	101.59	52.34	38	95.24	52.43	6.35	−17.76 to 30.46

The number of evaluable data points may not match those in Fig. 1 because of missing or incomplete questionnaires at specific time points.
(E3G, estrone-3-glucuronide).

found in Table III. The cortisol-to-creatinine ratio (and cortisol levels alone) was found to be generally higher in the control group over the test group (however, the data were highly correlated). There was a difference in variability at the different time points; the E3G data were more variable at surge times than Day 6 and concentrations were higher on the surge days in the test group compared with the control group. There appeared to be no relationship between total stress and endocrine levels throughout the study and no evidence of a relationship between pregnancy and endocrine levels in this sample of subjects.

Qualitative data analysis

A random selection of 30 (15 from each randomization group) exit interviews were used for analysis. The qualitative data collected
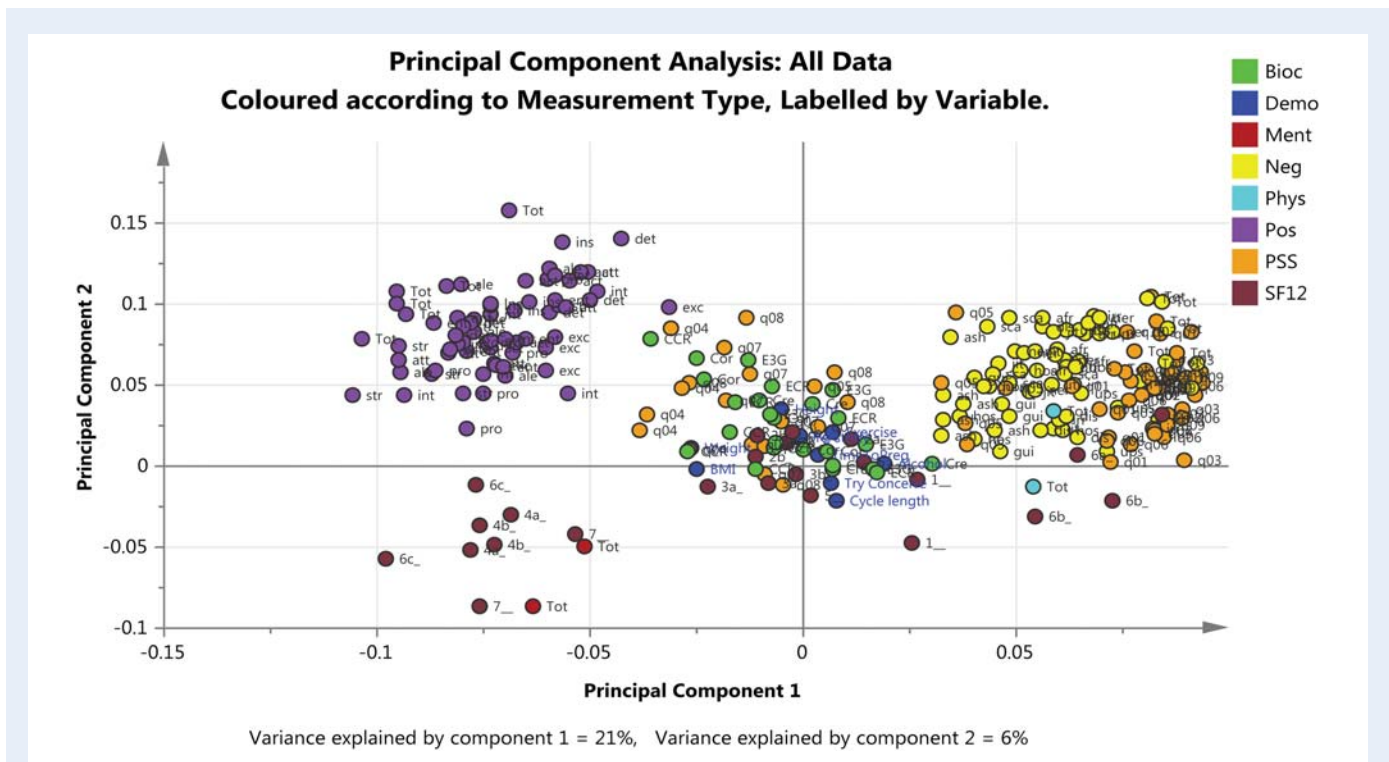
Figure 4 Loadings plot based on PCA fitted to all the data including all the questionnaires, biochemical markers and demographic characteristics. Principal component 1 (*x*-axis) relates to the trends observed between the positive and negative questions/responses from the PANAS and PSS questionnaires and principal component 2 (*y*-axis) is related to the positive and negative questions/responses from the SF-12 questionnaire. Two measurements that are close to each other are positively correlated and measurements that are diagonally opposite each other are negatively correlated. Clear separation is evident between the positive and negative responses to the PANAS questionnaire (purple and yellow dots, respectively) and to the PSS questionnaire as depicted by the two populations of orange dots observed. Bioc: Biochemistry marker, Demo: Demographic characteristic, Pos: Positive and Negative Affect Schedule (PANAS) Positive affect, Neg: PANAS: Negative affect, PSS: Perceived Stress Scale (PSS) Total Stress, Ment: Short Form 12 (SF12) Mental, Phys: SF12 Physical, SF12 Total score. On graph: Individual questions colour coded according to questionnaire, otherwise as indicated, Try Conceive: length in time trying to conceive (months), Pre-Study: length of pre-study period (days), BMI; Body Mass Index, Cre: Creatinine (g/dl), Cor: Cortisol (µg/dl), E3G: Estrone-3-Glucuronide (ng/ml), ECR: E3G:Creatinine ratio, CCR: Cortisol:Creatinine ratio.
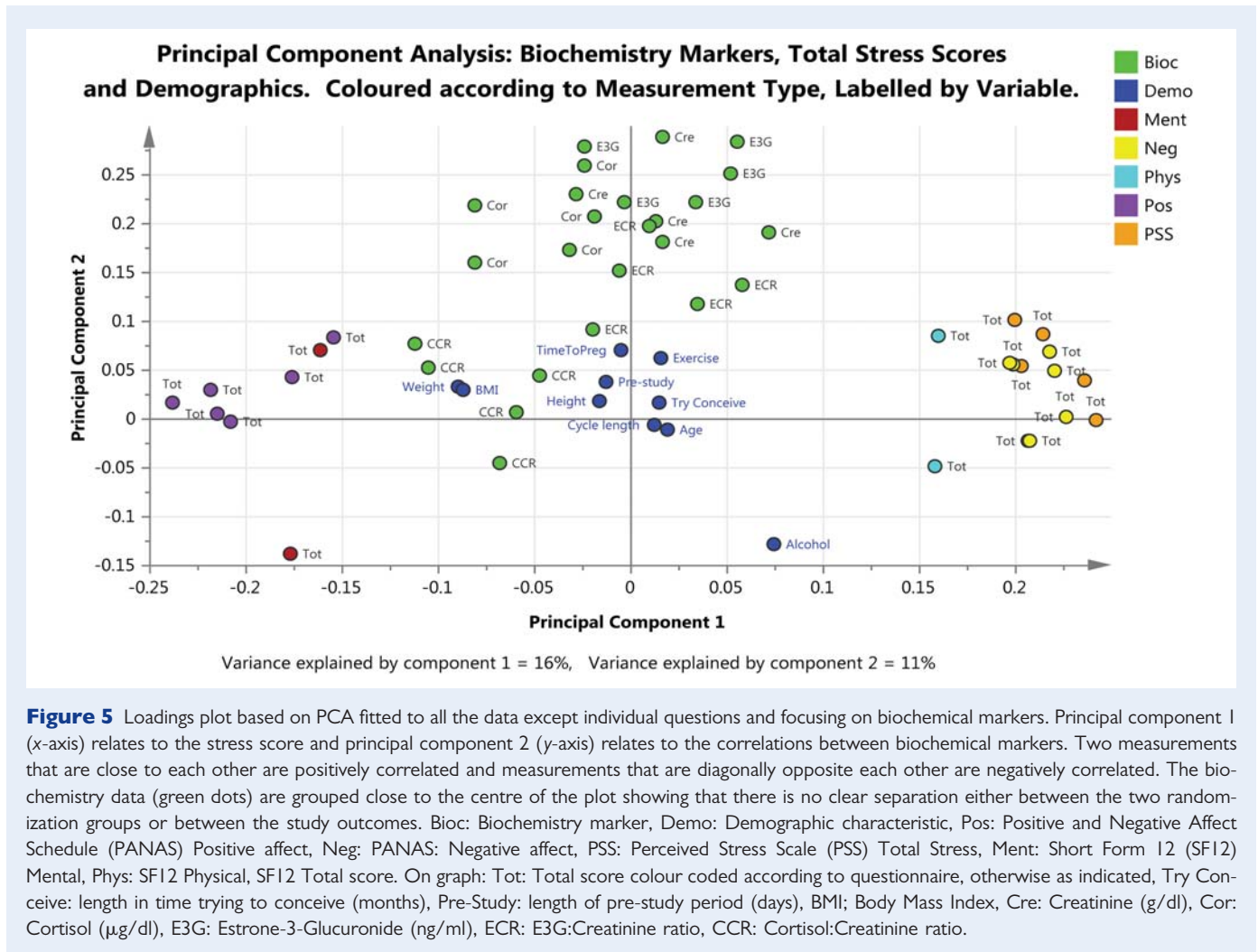
during the exit interview offered a deeper insight into the thoughts and feelings of the individuals who took part in the study. Results from closed questions in which the answers could be tabulated can be found in Table IV. These results highlight that many women expected to get pregnant within a short time frame, 56.7% stating that they would have expected to achieve a pregnancy either quickly or within 6 months of starting to try. It also highlighted that many women had little knowledge of the length of their cycles or when they were ovulating, 66.67% stating that before joining the study they had no idea of when they were ovulating. In addition, 53% of those analysed from the test group felt that they had a better understanding of changes occurring in their bodies around the time of ovulation after using the digital ovulation tests during the study.

Of the volunteers in the control group who were provided with the NICE guidelines for increasing chances of conception, 80% chose to follow them when asked during the exit interview. Those that chose not to follow these guidelines did so because of the perceived added pressure or inconvenience. All the interviewees found the digital ovulation tests easy to use and understand, often comparing them with the visual read tests, which they reported as being difficult to read.

Some of the perceived advantages and disadvantages of using the digital ovulation tests mentioned by volunteers during the exit interview have been listed and ordered by frequency of response in Table V. The most commonly mentioned advantage was the ability of the test to identify and pinpoint when ovulation occurs and this was linked by the volunteers to the increased likelihood of getting pregnant. It was also observed that using the ovulation tests helped the users to focus their conception attempts to the correct time in their cycle, one volunteer saying that she felt she would be doing something psychologically proactive. The main response when asked about disadvantages was that there were none; however, other disadvantages included added pressure on the partner and a lack of spontaneity associated with timing intercourse, although this was often countered by the desire to achieve a pregnancy.

Discussion

This study used three different tools to assess levels of stress in women attempting to conceive outside of any fertility investigations; validated questionnaire, urinary biomarkers and qualitative information collected by interview. We found no difference either in the

Figure 5 Loadings plot based on PCA fitted to all the data except individual questions and focusing on biochemical markers. Principal component 1 (x-axis) relates to the stress score and principal component 2 (y-axis) relates to the correlations between biochemical markers. Two measurements that are close to each other are positively correlated and measurements that are diagonally opposite each other are negatively correlated. The biochemistry data (green dots) are grouped close to the centre of the plot showing that there is no clear separation either between the two randomization groups or between the study outcomes. Bioc: Biochemistry marker, Demo: Demographic characteristic, Pos: Positive and Negative Affect Schedule (PANAS) Positive affect, Neg: PANAS: Negative affect, PSS: Perceived Stress Scale (PSS) Total Stress, Ment: Short Form 12 (SF12) Mental, Phys: SF12 Physical, SF12 Total score. On graph: Tot: Total score colour coded according to questionnaire, otherwise as indicated, Try Conceive: length in time trying to conceive (months), Pre-Study: length of pre-study period (days), BMI; Body Mass Index, Cre: Creatinine (g/dl), Cor: Cortisol (μg/dl), E3G: Estrone-3-Glucuronide (ng/ml), ECR: E3G:Creatinine ratio, CCR: Cortisol:Creatinine ratio.

questionnaire results or in the biomarker analysis between levels of stress in the women who were using the ovulation tests compared with the control group who were not using any methods to try and identify when they were ovulating. This indicates that use of the digital ovulation test in couples who are trying to conceive should not be dismissed with the presumption that they cause stress.

The questionnaire results for all tools over the duration of the study remained relatively flat and results were comparable with the published normative data in female populations. The similarity in the SF-12 data between the two groups at baseline and at the end of the study is particularly encouraging as it suggests that the general well-being of participants who did not become pregnant in both groups was similar and did not change over the duration of the study. The results of this study can be applied to the cohort of women recruited, i.e. those that have no known evidence of infertility and who are looking to conceive in a non-medical setting. A separate study would be needed to assess whether the same can be said for women who have been trying to conceive for many months and who may be attending a fertility clinic, as psychological stress in this group of women is likely to be quite different.

Emotional distress such as tension and worry is often portrayed as a contributing factor in failing to conceive quickly; however, the evidence

is somewhat inconclusive. This notion is often based on anecdotal evidence of natural conceptions after adoption or holidays. There is some evidence to suggest that psychological distress may be a risk factor for reduced fertility in women, particularly in those with long menstrual cycles (Takefman et al., 1990; Sanders and Bruce, 1997; Hjollund et al., 1999).

A recent study of women trying to conceive from the general population found that those with the highest quartile of salivary α-amylase concentration had slightly lower day-specific probabilities of conception but this was apparent only in the first cycle (Buck-Louis et al., 2011). A further publication, based on the same study population, showed no relationship between pregnancy and self-reported psychological stress as measured by questionnaires on Day 6 of the menstrual cycle (Lynch et al., 2012).

A recent meta-analysis by Boivin et al. (2011) of 14 prospective studies looking at the association between emotional distress and pregnancy in women undergoing a single cycle of assisted reproductive technology indicated a lack of association between emotional distress (e.g. feelings of tension, nervousness or worry) and pregnancy outcome. Other studies have also shown little or no effect of stress on pregnancy outcome (Milad et al., 1998; Sheiner et al., 2003; Anderheim et al., 2005). Our study results support the latter view as levels of

Table IV Qualitative data—responses to closed questions from the exit interview.

Question	Response	Test (%), (n = 15)	Control (%), (n = 15)
How long did you expect it would take you to conceive?	≤3 months	1 (7)	3 (20)
	≤6 months	4 (27)	2 (13.3)
	≤12 months	0 (0)	0 (0)
	≤24 months	1 (7)	2 (13.3)
	Do not know	4 (26)	6 (40)
	Not long/quickly	5 (33)	2 (13.3)
Before joining the study did you think you knew when you were ovulating?	Yes	4 (27)	6 (40)
	No	11 (73)	9 (60)
Have you ever used any method for determining when ovulation occurs?	Yes	7 (47)	8 (53)
	No	7 (47)	7 (47)
	Not asked	1 (7)	0 (0)
Do you feel any pressure or stress during your cycle?	Yes	2 (13)	10 (67)
	No	3 (20)	4 (26)
	Work/family bereavement	2 (13)	1 (7)
	Miscarriage	1 (7)	0 (0)
	Not asked	7 (47)	0 (0)
Does the prospect of testing for ovulation appeal to you?	Yes	15 (100)	15 (100)
	No	0 (0)	0 (0)
Did you follow the NICE guidelines for achieving conception while on the study?	Yes/aware	—	12 (80)
	No	—	3 (20)
Did the ovulation test meet your expectations when you used it?	Yes	14 (93)	—
	No	1 (7)	—
Was the product easy to use?	Yes	15 (100)	—
	No	0 (0)	—
Did using the test help you to notice any changes occurring in your body during your cycle?	Yes	8 (53)	—
	No	3 (20)	—
	Not asked	4 (27)	—
Did you feel any pressure or stress when using the test?	Yes (1 due to work/1 lack of test sticks)	7 (47)	—
	No	8 (53)	—
Will you consider purchasing this product in the future?	Yes	13 (86)	—
	No	1 (7)	—
	Possibly (as other arrangements with GP)	1 (7)	—

stress were found to be unrelated to pregnancy outcome, or indeed, affected by emotional factors such as number of previous miscarriages or length of time trying to conceive.

The potential effect of stress on pregnancy outcome is important as current guidelines suggest that ovulation tests can increase stress in women who are trying to conceive. The only study to date that specifically addresses this question is that by Kopitzke et al. (1991). Retrospective studies such as that used to support the NICE guidelines (Kopitzke et al., 1991), which measure stress when the outcome of the treatment is already known, may affect the participants' view of the level of stress they experienced (Boivin and Takefman, 1995). Prospective studies such as ours are likely to provide a more accurate appreciation of levels of stress in real time. Owing to the parity of data and plethora of opinion in this area, we believe that our study is extremely important as it provides new evidence that the use of digital ovulation tests by women outside of fertility clinics does not cause stress. Published evidence also suggests that there are likely to be far more stressful experiences while trying to conceive than timing of intercourse alone (Connolly et al., 1993; Boivin and Takefman, 1995; Boivin and Lancastle, 2010).

Although the number of pregnancies in this study was found to be higher in women who were using digital ovulation tests, the study was inadequately powered to compare the two groups with respect to conception. We can conclude that use of ovulation tests does not have a detrimental effect on time to conception. We suggest that a further study of more women over a longer period of time would be useful in assessing the effect of digital ovulation tests on conception rates. Much of the data available for ovulation tests currently focus on the accuracy of the tests rather than efficacy except for one study of a fertility monitor (which has an ovulation test component) reporting it to be associated with an increased pregnancy rate (Robinson et al., 2007). This adds weight to the argument that factors such as mistiming of intercourse, increasing age of women who are trying to conceive and a lack of knowledge of the menstrual cycle and cycle length are more likely than increased levels of stress, to affect the chances of conception in any given cycle (Robinson and Ellis, 2007; Zinaman et al., 2012).

Table V Summary of actual (test group) and perceived (control group) advantages and disadvantages of ovulation test use by randomization group.

Test group	Control group
Advantages	
Timing/pinpoint of ovulation	Timing/pinpoint ovulation
Ability to plan	Removes uncertainty
Easy to use	Easy to use
Relaxes/stress free/puts mind at ease	Know that you get a surge on a regular basis
Takes pressure off reading body/scientific	Very straight forward/clear
Learn about cycle	Convenience/keep focused
Improves chances of pregnancy	Psychologically proactive
Accuracy/very clear	Smiley face, not line
Know where in cycle if irregular	
Disadvantages	
None (33%)	None (33%)
Added pressure/pressure on partner	Knew optimum time so why not happening
Removes fun factor spontaneity/could get obsessed	Could get emotionally dependent
Having to test same time every day	Involvement/cannot take away on holiday/regimental
Distressing if calendar says ovulating but not	Costly. Likely to deter from intercourse
Gutted to see surge as partner was away	Process functional/pressure

Themes have been ordered by frequency of response.

For a number of years, women across Europe and other developed nations have been postponing parenthood (Mathews and Hamilton, 2009). According to recent population statistics in the UK, over the past 20 years (1990–2010), there has been a clear shift in the total number of live births to increased age of mother (ONS, 2010). It has also been shown that the monthly chance of achieving a clinical pregnancy or live birth starts declining in the late 20s for women, with a substantial decline by the late 30s (Schmidt et al., 2011). As a consequence, an increasing proportion of couples are experiencing infertility and prolonged time to pregnancy as well as a range of adverse pregnancy outcomes. These facts coupled with the modern-day desire and expectation that getting pregnant is something that should happen quickly and easily mean that more women are looking for tools that can help them achieve their goal. In our study, over half of the volunteers expected to conceive within 6 months of starting to try. However, because many women have a limited understanding of their own fertility (Small et al., 2007) and start their families later in life, getting pregnant can often take longer than expected.

The qualitative aspect of this study has highlighted a number of other positive benefits of using the digital ovulation test which for some women was very helpful. These include a sense of empowerment, and a feeling of comfort and reassurance that the test can show a surge in LH occurs and that intercourse can be focused to the right time of the cycle.

There may also be value in knowing early whether or not a surge in LH is occurring as it may be a trigger for seeking medical advice sooner. In many cases, particularly in older women, fertility advice is often not sought early enough. In fact, <55% of subfecund women have been shown to seek medical advice and treatment at all (Boivin et al., 2007). A study by Bunting and Boivin (2007) showed that often this is related to a fear of discovering a problem and these negative reactions can substantially delay seeking help in 20% of women. It is of interest that a proportion (14%) of the volunteers in our study had been trying to conceive for longer than 12 months as by definition, infertility is the inability to conceive after 12 months of unprotected intercourse (ASRM, 2006). This may be a reflection of the type of women in the UK population with possible fertility problems who do not seek medical advice when perhaps they should. Use of digital ovulation tests in the early attempts at conception may prompt women who do not detect the LH surge to seek help earlier than they otherwise would.

The variation in the E3G results on the surge day of both cycles was found to be larger in both the test and the control groups, probably because the collection of this sample was timed to the day of the LH surge, where estrogen concentration is also variable. In addition, the E3G: creatinine ratio was higher in the test group compared with the control group, probably because the collection of this sample was timed to the day of the LH surge, where urinary E3G is also maximal. For the control group, timing was assigned on the basis of self-reported average cycle length. This method provides much less accurate timing, so for many volunteers in the control group this did not correspond to the day of maximal urinary E3G. This shows that identification of peak fertility based on menstrual cycle characteristics and self-reported cycle length is not reliable.

A potential flaw in this study is that we saw an unexpectedly high pregnancy rate in the test group including a large number of pregnancies before the start of the study, and therefore, to reduce any weakness in the study population, we implemented an additional bias recruitment to ensure that there was a representative sample of test users who did not achieve a pregnancy at the end of the study. Although no significant differences were observed between the two phases of recruitment for any of the demographic information, there is a possibility that the second cohort of volunteers may have slightly different baseline characteristics. It is possible, for example, that the second recruitment phase may include those that were unsuccessful in the first phase and therefore may have been trying to conceive for a longer length of time.

This study did not observe the effect of use of the ovulation test on male partners. Several studies have shown that timed intercourse as required in early-fertility investigations for the post-coital test, for example, may increase emotional distress or cause sexual dysfunction in men (Drake and Grunert, 1979; Takefman et al., 1990; Boivin et al., 1992). Further studies of men examining the effect of timed intercourse using ovulation tests would be interesting using the same analysis tools used here.

A further research question would be whether or not use of home ovulation tests can influence the behaviour of couples in focusing intercourse to the fertile period by collecting data on timing and frequency

of intercourse. However, this type of information is difficult to collect accurately and can make participation over-cumbersome for the volunteers taking part with risk to overall study compliance and its inclusiveness to volunteers.

In summary, this study has shown that there is no difference in levels of stress between women using home ovulation tests and women who are trying to conceive having been provided with the NICE guidelines on increasing the chances of conception. The suggested benefits of use of home ovulation detection kits include possible reduction in time to pregnancy, improved understanding of ovarian and menstrual cycles and positive reinforcement of reproductive health. Women who wish to use these tests as an aid to conception should not be discouraged because of unfounded suggestions that they increase stress levels among users.

Acknowledgements

Authors' roles

S.J. and W.L. made substantial contribution to the concept and study design. S.T. contributed to execution of the study and acquisition of the data as well as drafting of the manuscript. G.J. and M.C. were involved with data analysis and interpretation. All the authors made substantial contributions to revising the article critically for important intellectual content and final approval of the version to be published.

Funding

Conflict of interest

S.T. and S.J. are employees of SPD Development Company Limited. W.L. is an independent clinician, who did not receive consultancy fees for his intellectual contribution to this study, but continues to provide paid consultancy to SPD

References

American Society for Reproductive Medicine (ASRM) Practice Committee. Definition of 'infertility'. *Fertil Steril* 2006;**86**(Suppl 4):S228.

Anderheim L, Holter H, Bergh C, Moller A. Does psychological stress affect the outcome of *in vitro* fertilization? *Hum Reprod* 2005;**20**:2969–2975.

Behre H, Kuhlage J, Gasner C, Sonntag B, Schem C, Schneider H, Nieschlag E. Prediction of ovulation by urinary hormone measurements with the home use ClearPlan Fertility Monitor: comparison with transvaginal ultrasound scans and serum hormone measurements. *Hum Reprod* 2000;**15**:2478–2482.

Blackwell LF, Brown JB, Vigil P, Gross B, Sufi S, d'Arcangues C. Hormonal monitoring of ovarian activity using the ovarian monitor, Part I. Validation of home and laboratory results obtained during ovulatory cycles by comparison with radioimmunoassay. *Steroids* 2003;**68**:465–476.

Boivin J, Lancastle D. Medical waiting periods: imminence, emotions and coping. *Women's Health* 2010;**6**:59–69.

Boivin J, Takefman J. Stress level across stages of in vitro fertilisation in subsequently pregnant and nonpregnant women. *Fertil Steril* 1995;**64**:802–810.

Boivin J, Takefman J, Brender W, Tulandi T. The effects of female sexual response in coitus on early reproductive processes. *J Behav Med* 1992;**15**:509–518.

Boivin J, Bunting L, Collins J, Nygren G. International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. *Hum Reprod* 2007;**22**:1506–1512.

Boivin J, Griffiths E, Venetis C. Emotional distress in infertile women and failure of assisted reproductive technologies: meta-analysis of prospective psychosocial studies. *Br Med J* 2011;**342**(d223):1–9.

Brown J, Blackwell L, Billings J, Conway B, Cox R, Garrett G, Holmes J, Smith M. Natural family planning. *Am J Obstet Gynecol* 1987;**157**:1082–1089.

Buck Louis G, Lum K, Sundaram R, Chen Z, Kim S, Lynch C, Schisterman E, Pyper C. Stress reduces conception probabilities across the fertile window: evidence in support of relaxation. *Fertil Steril* 2011;**95**:2184–2189.

Bunting L, Boivin J. Decision-making about seeking medical advice in an internet sample of women trying to get pregnant. *Hum Reprod* 2007;**22**:1662–1668.

Cohen S, Williamson G. Perceived stress in a probability sample of the United States. In: Spacapan S, Scamp S (eds). *The Social Psychology of Health: Claremont Symposium on Applied Social Psychology*. Newbury Park, CA: Sage Publications, 1998, 31–67.

Cohen S, Kamarch T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* 1983;**24**:385–396.

Collins W. Hormonal indices of ovulation and the fertile period. *Adv Contracept* 1985;**1**:279–294.

Connolly K, Edelmann R, Bartlett H, Cooke I, Lenton E, Pike S. An evaluation of counselling for couples undergoing treatment for *in-vitro* fertilisation. *Hum Reprod* 1993;**8**:1332–1338.

Crawford J, Henry J. The positive and negative affect schedule (PANAS): construct validity, measurement properties and normative data in large a large non-clinical sample. *Br J Clin Psychol* 2004;**43**:245–265.

Creinin M, Keverline SandMeyn L. How regular is regular? An analysis of menstrual cycle regularity. *Contraception* 2004;**70**:289–292.

Drake T, Grunert G. A cyclic pattern of sexual dysfunction in the fertility investigation. *Fertil Steril* 1979;**32**:542–545.

Gnoth C, Frank-Herrmann P, Freundl G. Opinion: natural family planning and the management of infertility. *Arch Gynecol Obstet* 2002;**267**:67–71.

Hilgers T, Daly K, Prebil A, Hilgers S. Cumulative pregnancy rates in patients with apparently normal fertility and fertility-focused intercourse. *J Reprod Med* 1992;**37**:864–866.

Hjollund N, Jenson T, Bonde J, Henriksen T, Andersson A, Kolstad H, Skakkebaek N, Olsen J. Distress and reduced fertility: a follow-up study of first-pregnancy planners. *Fertil Steril* 1999;**72**:47–53.

Jenkinson C, Stewart-Brown S, Peterson S, Paice C. Assessment of the SF-36 version 2 in the United Kingdom. *J Epidemiol Community Health* 1999;**53**:46–50.

Johnson S, Miro F, Barrett S, Ellis J. Levels of urinary human chorionic gonadotrophin (hCG) following conception and variability of menstrual cycle length in a cohort of women attempting to conceive. *Curr Med Res Opin* 2009;**25**:741–748.

Johnson S, Ellis J, Godbert S, Ali S, Zinaman M. Comparison of a digital ovulation test with three popular line ovulation tests to investigate user accuracy and certainty. *Expert Opin Med Diagn* 2011;**5**:467–473.

Kopitzke E, Berg B, Wilson J, Owens D. Physical and emotional stress associated with components of the infertility investigation: perspectives of professionals and patients. *Fertil Steril* 1991;**55**:1137–1143.

Lokuge S, Frey B, Foster J, Soares C, Steiner M. Depression in women: windows of vulnerability and new insights into the link between estrogen and serotonin. *J Clin Psychiatry* 2011;**72**: e1563–e1569.

Lynch C, Sundaram R, Buck-Louis G, Lum K, Pyper C. Are increased levels of self-reported psychosocial stress, anxiety, and depression associated with fecundity? *Fertil Steril* 2012;**98**:453–458.

Marshall C, Rossman G. *Designing Qualitative Research*, 3rd edn. London: Sage, UK, 1999.

Mathews T, Hamilton B. 2009. Delayed childbearing: more women are having their first child later in life. NCHS Data Brief 21 [online] Available at: <http://www.cdc.gov/nchs/data/databriefs/db21.pdf> (10 August 2012, date last accessed).

Matthews JNS, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. *Br Med J* 1990;**300**:230–235.

Miki K, Sudo A. Effect of urine pH, storage time and temperature on stability of catecholamines, cortisol and creatinine. *Clin Chem* 1998; **44**:1759–1762.

Milad M, Klock S, Moses S, Chatterton R. Stress and anxiety do not result in pregnancy wastage. *Hum Reprod* 1998;**13**:2296–2300.

Miro F, Parker S, Aspinall L, Coley J, Perry P, Ellis J. Sequential classification of endocrine stages during reproductive aging in women: the FREEDOM study. *Menopause* 2005;**12**:281–290.

Nakamura K, Sheps S, Arck P. Stress and reproductive failure: past notions, present insights and future directions. *J Assist Reprod Genet* 2008;**25**:47–62.

National Institute for Clinical Excellence (NICE) Clinical Guideline, 2004. Fertility: assessment and treatment for people with fertility problems. [online] Available at: <http://publications.nice.org.uk/fertility-cg11> (13 March 2012, date last accessed).

Nepomnaschy P, Altman R, Watterson R, Co C, McConnell D, England B. Is cortisol excretion independent of menstrual cycle day? A longitudinal evaluation of first morning urinary specimens. *PLoS ONE* 2011;**6**:e18242.

Nicolson N. Measurement of cortisol. In: Luecken L, Gallo L (eds). *Handbook of Physiological Research Methods in Health Psychology*. Thousand Oaks, CA: Sage Publications, 2008, 37–74.

Office for National Statistics (ONS). 2010. Live births in England and Wales by characteristics of mother 2010. [online] Available at: <http://www.ons.gov.uk/ons/index.html> (13 March 2012, date last accessed).

Robinson J, Ellis J. 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**:301–306.

Robinson J, Wakelin M, Ellis J. Increased pregnancy rate with use of the Clearblue Easy Fertility Monitor. *Fertil Steril* 2007;**87**:329–334.

Sanders K, Bruce N. A prospective study of psychosocial stress and fertility in women. *Hum Reprod* 1997;**12**:2324–2329.

Schmidt L, Sobotka T, Bentzen J, Andersen A. Demographic and medical consequences of the postponement of parenthood. *Hum Reprod Update* 2011;**18**:29–43.

Severy L, Robinson J, Findley-Klein C, McNulty J. Acceptability of a home monitor used to aid in conception: psychosocial factors and couple dynamics. *Contraception* 2006;**73**:65–71.

Sheiner E, Sheiner E, Potashnik G, Carel R, Shoham-Vardi I. The relationship between occupational psychological stress and female fertility. *Occup Med* 2003;**53**:265–269.

Singh M, Baxena B, Rathnam P. Clinical validation of enzyme immunoassay of human luteinizing hormone (hLH) in the detection of the pre-ovulatory luteinizing hormone (LH) surge in urine. *Fertil Steril* 1984;**41**:210–217.

Small C, Manatunga A, Marcus M. Validity of self-reported menstrual cycle length. *Ann Epidemiol* 2007;**17**:163–170.

Stanford J, White G, Hatasaka H. Timing intercourse to achieve pregnancy: current evidence. *Obstet Gynecol* 2002;**100**:1333–1341.

Takefman J, Brender W, Boivin J, Tulandi T. Sexual and emotional adjustment of couples undergoing infertility investigation and the effectiveness of preparatory information. *J Psychosom Obstet Gynecol* 1990;**11**:275–290.

Ware J, Sherbourne C. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; **30**:473–483.

Watson D, Clark L, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol* 1988;**54**:1063–1070.

WHO Task force on Methods for Determination of the Fertile Period. Temporal relationships between ovulation and defined changes in the concentration of plasma estradiol-17β, luteinising hormone, follicle stimulating hormone and progesterone. I. Probit analysis. *Am J Obstet Gynaecol* 1980;**138**:383–390.

Wilcox A, Dunson D, Baird D. The timing of the 'fertile window' in the menstrual cycle: day specific estimates from a prospective study. *Br Med J* 2000;**321**:1259–1262.

Zinaman M, Johnson S, Ellis J, Ledger W. Accuracy of perception of ovulation day in women trying to conceive. *Curr Med Res Opin* 2012;**28**:1–6.