Comparison of three infertility-specific measures in men and women going through ART treatment

Running title: comparing infertility-specific measures in ART patients

Juliana Pedro[1], Yoon Frederiksen[2], Lone Schmidt[3], Hans J Ingerslev[4], Robert Zachariae[2,4] and Mariana V Martins[1]

Juliana Pedro
(Faculty of Psychology and Education Sciences) University of Porto, PT

Yoon Frederiksen
(Department of Psychology and Social Sciences) Aarhus University, DK

Lone Schmidt
(Social Medicine Department) University of Copenhagen, DK

Hans J Ingerslev
(Center for Preimplantation Genetic Diagnosis/The fertility Clinic) Aarhus University Hospital, DK

Robert Zachariae
(Department of Oncology Aarhus) University Hospital & (Department of Oncology) Aarhus University, DK

Mariana V Martins
(Faculty of Psychology and Education Sciences) University of Porto, PT

Corresponding author: Mariana V Martins, Faculty of Psychology and Education Sciences, University of Porto, 4200 Porto, Portugal.
Email: mmartins@fpce.up.pt

Abstract

We compared the psychometric properties of COMPI Fertility Problem Stress Scales (COMPI-FPSS), Fertility Problem Inventory (FPI) and Fertility Quality of Life Tool (FertiQoL) in 293 patients enrolled for assisted reproductive technology. COMPI-FPSS and FPI subscales presented higher internal consistency. COMPI-FPSS differentiated best between its domains. FPI revealed better concurrent validity. FertiQoL presented better fit. While discrimination for depression was similar between measures, FertiQoL was better at discriminating anxiety. Results suggest that while all compared measures are reliable and valid in assessing the psychosocial adjustment to infertility, the choice of measure should be based according to the assessment goals.

Keywords:

Infertility, psychosocial adjustment, infertility-specific measures, psychometrics, reliability and validity

Introduction

Infertility affects about 9% of couples in reproductive age (Boivin et al., 2007).
The impact of an infertility diagnosis has been associated with high emotional burden
(Luk and Loke, 2015), which can persist even when a pregnancy is achieved (HaCohen
et al., 2016). Evidence showed that ≅34%-47% of fertility treatment patients are at risk
of maladjustment (Lopes et al., 2014; Verhaak et al., 2010), who sometimes feel
embarrassed seeking help and support (Hanna and Gough, 2016). While measures of
general psychological adjustment (e.g., anxiety, depression) have been used initially to
assess the impact associated with this challenge (Emery et al., 2003; Salvatore et al.,
2001; Slade et al., 1997; Smeenk et al., 2001), some authors claimed that the use of
general psychological instruments may not capture the specific experience of infertility
(Boivin, 2003; Greil et al., 2016; Schmidt, 2009). This has stimulated the development
of more than 10 self-report measures assessing infertility psychosocial adjustment (see
reviews of Gourounti et al., 2010; Martins et al., 2016), the most used being the COMPI
Fertility Problem Stress Scales (COMPI-FPSS, Schmidt, 2006; Schmidt, Holstein, et al.,
2005), the Fertility Problem Inventory (FPI, Newton et al., 1999) and the Fertility
Quality of Life Tool (FertiQoL, Boivin et al., 2011).

These three measures were recently identified as important tools in screening for
fertility treatment adjustment (ESHRE Psychology and Counseling Guideline
Development Group, 2015). Research using these measures has provided significant
contributions to the literature concerning gender differences (Chachamovich et al.,
2009; Peterson et al., 2003; Peterson, CR Newton, et al., 2006; Schmidt et al., 2003;
Slade et al., 2007), the predicting roles of coping strategies and social support (Keramat
et al., 2014; Martins et al., 2011, 2013; Peterson, C Newton, et al., 2006; Schmidt,

Holstein, et al., 2005) and the effect infertility-stress on marital satisfaction (Gana and Jakubowska, 2016).

Investigating the psychometric properties of instruments to measure infertility-specific adjustment has been considered a research priority (Greil, 1997; Schmidt, 2009), the main goal being to determine clinical usefulness (e.g., identifying clinical cutoffs). This would help clinicians to correctly detect patients who are vulnerable to maladjustment and emotional problems, and refer them to counseling or psychotherapy, with targeted priority domains of intervention (Boivin et al., 2011; ESHRE Psychology and Counseling Guideline Development Group, 2015; Verhaak et al., 2010). Given that the literature has shown that an optimal adaptation to the emotional consequences of infertility and its treatments might contribute to lower drop-out (e.g, Eisenberg et al., 2010; Verberg et al., 2008), and better treatment outcomes (Klonoff-Cohen et al., 2001; Lancastle and Boivin, 2005), this early detection is crucial. Clinical usefulness requires the examination of reliability, validity, factor structure (Byrne, 1998; Streiner and Norman, 1995, 2008), and determination of cutoffs. Although some of these properties have previously been investigated separately for all three measures (Aarts et al., 2011; Boivin et al., 2011; Donarelli et al., 2016; Hsu et al., 2013; Melo et al., 2012; Moura-Ramos et al., 2012), their psychometric properties and ability to discriminate clinically significant emotional problems have not yet been compared in the same sample.

The present study explores the COMPI-FPSS, the FertiQoL and the FPI factorial structure, convergent and concurrent validity, and reliabilities in a cross-sectional sample of patients seeking fertility treatment. Furthermore, the ability of these three measures to discriminate clinically relevant depression and anxiety was tested. By eliminating differences in sample characteristics and conducting thorough statistical analyses for all measures, it is hoped that future researchers and clinicians will have

accurate information concerning the properties of the three measures on which to base their selection.

Methods

Participants and Procedure

From November 2010 through July 2012, 651 women and their partners (*n*=1302) attending three Danish fertility clinics received a mailed invitation to participate and 383 agreed. Then, 90 were excluded for not meeting the following eligibility criteria: a) enrollment for in-vitro fertilization (IVF) or intra-cytoplasmatic sperm injection (ICSI), b) being married or partnered, c) being between 18 and 45 years old, c) not previously diagnosed with pre-implantation genetic diagnosis d) ability to understand and write Danish. The final sample had 293 subjects (161 women, 132 men). Participants completed all questionnaires at the same time point. Participants had a marriage/cohabitation length of 5.1 years (*SD*=1.10), and 52.2% had already undergone 1 cycle of IVF/ICSI. On average, patients had been trying to conceive for 2.52 years (*SD*=1.20). No statistically significant differences in reproductive characteristics were observed between those who accepted to participate in the study and non-responders (data not shown).

This study was approved by the Danish Regional Ethical Committee and the Danish Data Protection Agency. Permission to use the instruments was obtained from the authors of the FPI and COMPI-FPSS. The FertiQoL is freely available at its website.

Measures

Socio-demographical data and reproductive characteristics were obtained through questionnaires and medical records.

The *COMPI Fertility Problem Stress Scales* (Schmidt, 2006; COMPI-FPSS, Schmidt, Holstein, et al., 2005) was originally developed to the Danish population and includes 14 items assessing the impact of infertility on Personal, Social, and Marital domains (e.g., My life has been disrupted because of this fertility problem). Items are based on four or five-point Likert scales, with higher scores indicating higher stress. Internal consistencies are shown in Supplementary file, Table 1.

The *Fertility Problem Inventory* (FPI, Newton et al., 1999) assesses infertility-related stress, producing a total and five subscale scores: Relationship concern, Social concern, Sexual concern, Rejection of a child-free lifestyle and Need for parenthood. The 46 FPI items (e.g., I feel empty because of our fertility problem) are scored on a 6-point Likert scale, with higher scores indicating higher infertility-related stress. The FPI was translated into Danish and back-translated by an English native speaker, leading to minor adjustments. Cronbach's alpha values are shown in Supplementary file, Table 2.

The *Fertility Quality of Life Tool* (FertiQoL; Boivin et al., 2011) was developed to assess QoL in women and men experiencing fertility problems, and is available in several languages including Danish (see Fertility Quality of Life tool, 2016 retrieved from http://sites.cardiff.ac.uk/fertiqol/download/). The measure is organized in two modules: Core FertiQoL and Optional Treatment FertiQoL. The Core FertiQoL includes 4 subscales: Emotional, Mind/body, Relational, and Social (e.g., Do you feel sad and depressed about your fertility problems?). The Optional Treatment FertiQoL module has 2 subscales: Environment and treatment Tolerability. Each question is rated on a 4-point Likert scale with higher scores indicating more favorable QoL. Internal consistencies are shown in Supplementary file, Table 3.

Depressive symptoms were assessed with the *Beck Depression Inventory* II (Beck et al., 1996; Christensen et al., 2009; Zachariae et al., 2004) which is composed of 21

items scored between 0 and 3. Higher scores indicate higher depressive symptoms. Internal consistency was .91 in the present sample.

State anxiety was assessed through the *State Trait Anxiety Inventory - Form Y* (Spielberger et al., 1983; Zachariae et al., 2001). This subscale has 20 items measuring anxiety at the moment of response in a 4-point Likert scale, with higher scores indicating higher state anxiety symptoms. Internal consistency was .94 in the present sample.

Statistical analysis

Descriptive statistics, reliability indices and correlations were calculated using IBM SPSS, Version 19. Reliability was measured through internal consistency (values ≥.70 considered acceptable (Streiner and Norman, 2008) and individual item analyses (reliability if item deleted). Convergent and concurrent validity were analyzed through correlations between the three infertility-specific measures and between each scale and generalized psychological adjustment measures BDI-II and STAI-State. Correlations were considered strong if ≥.70; moderate between .40 and .69; and weak ≤.39. Overlap of constructs was considered if r>.85 (Kenny, 1979).

Factor validity was evaluated through confirmatory factor analysis (CFA) using AMOS 19 with maximum likelihood estimation. In contrast to exploratory factor analysis, CFA is more appropriate for testing established theories and minimizing the role of chance (Byrne, 1998). Patients with missing values were excluded from the respective CFAs (one from COMPI-FPSS, six from FPI, one from FertiQoL). To improve structural validity, items were removed if elimination resulted in substantial increase in the subscale alpha, factor loading was below .40, or correlations with other factors were high (Kahn, 2006). Three comparative fit indices were appraised to test goodness of fit: X^2 statistics, Comparative Fix Index (CFI) and Root Mean Square Error

of Approximation (RMSEA). We followed Hooper et al.´s (Hooper et al., 2008)

guidelines for goodness of fit: χ^2/d.f. <5; RMSEA≤.07 , and CFI≥.95.

A receiver-operator characteristic (ROC) analysis using MedCalc (version 16.2)

was performed to assess discriminative validity, with anxiety and depression as

categorical variables (cutoff values: BDI>13 (Beck et al., 1996); STAI-state ≥40

(Cochrane et al., 2011; Frasure-Smith et al., 1995)). To conduct the ROC analysis,

FertiQoL was recoded to yield scores in the same direction as in the two remaining

measures. First, the Youden index (calculated by the formula sensitivity + specificity -1,

(Youden, 1950)) was used to determine each measure cut-off score. The Area Under the

Curves (AUC) was used to evaluate the accuracy of the three measures (excellent:>.90,

good:.90-.80, fair:.70-.79 and poor:.60-.69 (Hosmer and Lemeshow, 2000)). This area

represents the probability of correctly distinguishing a randomly chosen case from a

randomly chosen non-case based on the screening scales (Kessler et al., 2002). Then,

pairwise comparisons were performed to compare the overall performance of the three

measures in distinguishing patients with and without clinically relevant depression and

state anxiety.

Results

Descriptive statistics

Table 1 presents mean values, standard deviations and inter-correlations among

the three measures. Correlations between the COMPI-FPSS subscales were moderate.

In the FPI, Sexual and Relationship concerns were moderately inter-correlated, but were

both weakly correlated with Rejection of Childfree-lifestyle and Need for Parenthood,

which in turn were strongly inter-correlated. Moderate correlations were found between

Social concern and the remaining four subscales. Correlations between FertiQoL

subscales showed more variation, with strong correlations between the FertiQoL

Emotional, Mind/body and Social subscales and weak correlations between the Relational subscale and remaining subscales. While the subscale assessing Environment was also weakly correlated with the other subscales, the Tolerability subscale was strongly associated to the Emotional and Mind/Body subscales, and weakly to the Relationship and Environment subscales, with a moderate association observed only with the Social subscale.

[Insert Table 1 here]

Reliability

Estimates for the total scale and subscales of the COMPI-FPSS were acceptable to good (Cronbach's alphas between .73 and .89, Supplementary file, Table 1). Eight items were removed from the FPI and three items were removed from the FertiQoL due to individual item reliability analysis (see Supplementary file, Table 2 and 3). After removing these items and six more with factor loadings below .40, internal consistency increased in both measures, with estimates for the FPI total scale and subscales now being good to excellent (.81 to .93). The FertiQoL revealed excellent internal consistency for the total score (.91), but some subscales exhibited lower values (e.g., FertiQoL Tolerability, $\alpha=.54$).

Factor Structure

A first-order model and a second-order model were tested for each measure. Additionally, an alternative factor structure proposed by Moura-Ramos, Gameiro and Canavarro (2012) was tested for the FPI (Supplementary file, Table 4). We retained the FertiQoL authors' theoretical original structure for the second-order model, with four main and two optional factors separately loading on a 'core' factor and a 'treatment' factor, respectively, and these two latent variables loading on a general factor.

Analyses of the COMPI-FPSS revealed that both first-order and second-order models improved after adding two paths to show covariance between the error terms of the Personal subscale. The first-order model CFA yielded acceptable fit indices, while the second-order model showed poor fit. The chi-square difference test confirmed that the first-order model had significantly better fit ($\Delta\chi^2$=42.59; Δdf =1; $p<.001$). Standardized factor loadings of the 14 items (Supplementary file, Table 1) and standardized correlations among the three factors were statistically significant ($p<.001$).

Eight FPI items were removed and substantially increased alpha values in both models (Supplementary file, Table 2). By performing random parceling to decrease models' complexity (Little et al., 2002), model fit improved for all models. The first-order model showed a better fit than the second-order model ($\Delta\chi^2$=73; Δdf =5; $p<.001$) and Moura-Ramos et al.'s (2012) proposal, ($\Delta\chi^2$=11.61; Δdf =4; $p<.05$). Standardized factor loadings of each of the 13 parcels in the final model and standardized correlations between the five factors were all significant ($p<.001$).

FertiQoL had nine items removed (see Supplementary file, Table 3). The final models had 25 items, and we retained the authors' theoretical original structure for the second-order model, with four main and two optional factors separately loading on a 'core' factor and a 'treatment' factor, respectively, and these two latent variables loading on a general factor both showed acceptable fit indices, but the higher-order model showing poorer fit than the first-order model ($\Delta\chi^2$=43.32; Δdf=8; $p<.001$). Standardized factor loadings and correlations between the six factors were statistically significant ($p<.001$), except for associations between the Environment subscale with the Emotional, Mind/body, Social, and Tolerability subscales.

Convergent and concurrent validity

The three infertility-specific measures were highly correlated (p<.001), which indicates good convergent validity between COMPI-FPSS and FPI: .76; FPI and FertiQoL: -.76; COMPI-FPSS and FertiQoL: -.83.

Concurrent validity for COMPI-FPSS and FPI was supported by moderated correlations in the expected direction with both depression (.66 and .65, respectively) and state-anxiety (.60 and .59). Stronger correlations were found between FertiQoL total scores and both depression (-.76) and anxiety (-.73).

Discriminative validity

The Youden index indicated that the best sensitivity and specificity cut-off score points in distinguishing clinically relevant depression were 18 for the COMPI-FPSS (Youden =.55; sensitivity 81%; specificity 73%), 116 for the FPI (Youden =.54; sensitivity 88%; specificity 66%), and 71 for the FertiQoL (Youden =.60; sensitivity 87%; specificity 72%). In discriminating clinically relevant depression, all three scales revealed good AUC indexes (COMPI-FPSS: .85 (95%CI:.80-.89); FPI: .84 (95%CI:.79-.88); FertiQoL: .87 (95%CI:.83-.91)). The statistical pairwise comparison of diagnostic AUCs showed no differences in the performance between the three measures discriminating depression (p>.05).

In distinguishing patients with and without clinically relevant anxiety, the Youden index indicated the best sensitivity and specificity cut-off score points were 15 for COMPI-FPSS (Youden index =.37; sensitivity 75%; specificity 61%), 124 for FPI (Youden index =.44; sensitivity .71%; specificity 74%) and 71 for FertiQoL (Youden index=.51, sensitivity 78%; specificity 73%). The AUC was .75 (95%CI:.69-.80) for the COMPI-FPSS, .76 (95%CI:.71-.81) for the FPI and .82 (95%CI:.78-.87) for the FertiQoL. Pairwise comparison of ROC curves revealed that FertiQoL has significantly

better diagnostic performance than COMPI-FPSS (z=4.07; *p*<.001) and FPI (z=2.73; *p*=.006). The performance of COMPI-FPSS and FPI were similar (*p*=.63).

Discussion

This was the first study to evaluate and compare the psychometric properties of three infertility-specific measures: COMPI-FPSS, FPI, and FertiQoL.

All three measures and respective subscales presented good reliability indexes given the proposed factorial structures. The FertiQoL Relational and Tolerability subscales were the exceptions, but low alphas can be attributed to removal of items to increase reliability and validity (Briggs and Cheek, 1986; Pallant, 2010). These results are consistent with previous studies showing acceptable to good internal consistency for the FertiQoL subscales (Aarts et al., 2011; Boivin et al., 2011; Melo et al., 2012), and good consistency for COMPI-FPSS (Schmidt et al., 2003; Schmidt, Christensen, et al., 2005; Schmidt, Holstein, et al., 2005) and FPI (Martins et al., 2011; Moura-Ramos et al., 2012; Newton et al., 1999).

While the factorial structure of COMPI-FPSS was confirmed with acceptable fit, several modifications were necessary for FPI and FertiQoL to achieve acceptable fit. Given that COMPI-FPSS was constructed in a Danish population and based on input from Danish couples, it is possible that the items removed in FertiQoL and FPI are related to cultural validation issues, suggesting that certain items may not be generalizable across cultures. For example, the removal of the items "Family members don´t seem to treat us any differently" (FPI) and "Do you feel social pressure on you to have children?" (FertiQoL) might be due to a low social pressure to have children in Denmark compared to other cultures. Besides the fact that the majority of Danish patients discloses the fertility problem to their support network (Schmidt, Holstein, et

al., 2005), other social circumstances that reduce the social pressure to have children might be the high expectations for adults to take part in the labor force irrespective of parenthood, and the high rate of childless fifty-year old men (12,5%) and women (21%) (Statistics Denmark, 2015).

The FertiQoL factorial structure revealed excellent goodness of fit. However, the fact that the optional Tolerability subscale was only weakly correlated with the optional Environment subscale but strongly correlated with the Emotional and the Mind/body subscales suggests that some constructs might be not as well defined as others. As for the FPI, even though the strong association between Rejection of childfree-lifestyle and Need for parenthood could support the inclusion of a higher-order 'representation about the importance of parenthood' factor suggested by Moura-Ramos et al. (2012), and Donarelli et al. (2015), the second-order model did not show good fit. Further research exploring alternative factor structures while comparing different cultural samples is needed to clarify these relationships.

Significant and strong associations between the three infertility-specific measures supported convergent validity. Although there was no overlap of constructs following the .85 threshold (Kenny, 1979), high associations between COMPI-FPSS and FertiQoL suggest that there could be some overlap between the constructs of infertility-related stress and fertility QoL. The high correlations between COMPI-FPSS Personal stress and FertiQoL Emotional, Mind/body, Social and FPI Social concern highlight the common aspects evaluated by all three measures, such as the physical, emotional and social consequences of facing fertility problems. Concurrent validity was also found in all three scales, with results revealing FPI and COMPI-FPSS as more discriminant measures than FertiQoL comparing with depression and state anxiety. Nevertheless, some studies have found only moderate correlations between FertiQoL and other

instruments of general distress, such as the Brief Symptom Inventory (BSI; Derogatis, 2000) (Melo et al., 2012) and the HADS (Aarts et al., 2011). The three measures were similarly good in discriminating patients with and without clinically relevant depression. Concerning the ability to discriminate infertile patients with or without clinically relevant state anxiety, FertiQoL was the only measure showing good performance, with the remaining two showing fair performance (AUCs below .80). The FertiQoL was able to correctly identified 79% of patients with clinical state anxiety and 73% of non-anxious patients.

When choosing which instrument to use in a given research or clinical setting, psychometric results found in the present study should be taken into consideration, with the choice of the appropriate measure depending on the clinician or researcher's goals. The COMPI-FPSS offers a practical advantage compared with FPI and FertiQoL because it has considerably less items and has three interrelated but independent domains of infertility-related stress. It is a quick and easy to complete self-report measure which patients can fill throughout demanding treatment stages, putting minimal burden on the patients. However, other areas affected by fertility problems are not covered by this measure. If the goal of a given scholar or fertility counselor is to assess long-term infertile patients' decision-making, FPI might be a better choice considering convergent validity findings. Given that women with persistent desire for parenthood seem to experience poor psychological adjustment five years after unsuccessful treatment (ESHRE Psychology and Counseling Guideline Development Group, 2015), Rejection of childfree-lifestyle and Need for parenthood subscales can be useful to assess patients' perceptions on treatment ending and alternative options to parenthood. If the interest lies in investigating how patients tolerate treatment, the optional FertiQoL Environment and Tolerability subscales are unique in assessing this aspect.

Additionally, all three measures seem good at discriminating patients with and without clinically relevant depression. However the FertiQoL was the best in discriminating anxiety and might be useful as a screening tool to identify patients in need for psychological support.

This study has some strengths that should be highlighted. To the best of our knowledge, this is the first work to extensively compare three fertility-specific psychosocial measures in a sample of patients seeking fertility treatment. Besides advancing knowledge on infertility-related measures psychometric properties, our findings revealed cut-off points to identify patients with clinically relevant depression and state anxiety. Both the sample size and the statistical procedures could assure reliability of results.

Limitations also need to be addressed. First, these self-reported measures were analyzed in a sample of patients seeking treatment who had already received an infertility diagnosis. Perceptions on the psychosocial impact of fertility problems may be different for couples who are not seeking treatment (e.g., worry about being diagnosed as infertile, Bunting and Boivin, 2007), and the diagnosis can either appease or raise more concerns on a couple (e.g., achieving the desired result, Klonoff-Cohen and Natarajan, 2004). Second, the results of the present study could be culturally biased. The COMPI-FPSS was developed in a Danish population. This could, for example, explain the better differentiation between subscales in this measure, compared to the remaining two (translated into Danish). It should be noted, however, that seven of the 16 items of COMPI-FPSS were adapted from the Fertility Problem Inventory developed in a US context (Abbey et al., 1991). Third, as Cronbach's alpha is highly dependent on the number of items (Briggs and Cheek, 1986; Pallant, 2010), the between-measures comparability of internal consistencies could have been affected. However, it could be

noted here that the similarity of some items formulations in COMPI-FPSS could also introduce some bias in the alphas scores. Fourth, because assessment was not repeated, test-retest reliability and responsiveness were not evaluated. Since evidence suggest that infertility-stress increases during the waiting phase of a pregnancy test (Boivin et al., 1998), this is an important point in the assessment of validity of fertility-specific adjustment measures. Finally, this study had a significant attrition rate, which could be explained by the fact that the invitation letter was mailed only women, and hence if they decided to ignore it the chance that their partners would receive the invitation or reply becomes low. Furthermore, because infertility is a shared stressor, it would be important to study the psychometric characteristics of these three measures using a dyadic strategy analysis.

Future studies looking at the psychometric properties of these measures should therefore attempt to assess responsiveness and predictive validity to compare the capability of these measures in detecting changes in fertility-specific adjustment over time, and in predicting important consequences such as depression after a failed cycle or post-partum depression. Further research is also needed to test the cross-cultural validity and investigate the patient acceptable symptom state (PASS) for these three measures. The PASS gives clinicians and researchers a cut-off representing the highest level of symptomatology in which patients consider themselves well (Kvien et al., 2007), which would allow clinicians to better discriminate if the strain reported by a patient in a particular questionnaire is indicative of a need for psychological support.

Declaration of Conflicting Interests

The Authors declare that there is no conflict of interest.

Funding

References

Aarts JWM, VanEmpel IWH, Boivin J, et al. (2011) Relationship between quality of life and distress in infertility: A validation study of the Dutch FertiQoL. *Human Reproduction* 26: 1112–1118.

Beck AT, Steer RA and Brown GK (1996) *Beck Depression Inventory: Manual*. 2nd ed. Sant Antonio: The Psychological Corporation.

Boivin J (2003) A review of psychosocial interventions in infertility. *Social Science and Medicine* 57: 2325–2341.

Boivin J, Andersson L, Skoog-Svanberg A, et al. (1998) Psychological reactions during in-vitro fertilization: similar response pattern in husbands and wives. *Human reproduction* 13: 3262–3267.

Boivin J, Bunting L, Collins J, et al. (2007) International estimates of infertility prevalence and treatment-seeking: Potential need and demand for infertility medical care. *Human Reproduction* 22(6): 1506–1512.

Boivin J, Takefman J and Braverman A (2011) The fertility quality of life (FertiQoL) tool: Development and general psychometric properties. *Human Reproduction* 26: 2084–2091.

Briggs SR and Cheek J (1986) The role of factor analysis in the development and evaluatbn of personality scales. *Journal of Personality* 54(1): 106–148.

Bunting L and Boivin J (2007) Decision-making about seeking medical advice in an internet sample of women trying to get pregnant. *Human Reproduction* 22: 1662–1668.

Byrne B (1998) *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chachamovich J, Chachamovich E, Fleck MP, et al. (2009) Congruence of quality of life among infertile men and women: Findings from a couple-based study. *Human Reproduction* 24: 2151–2157.

Christensen S, Zachariae R, Jensen AB, et al. (2009) Prevalence and risk of depressive symptoms 3-4 months post-surgery in a nationwide cohort study of Danish women treated for early stage breast-cancer. *Breast Cancer Research and Treatment* 113: 339–355.

Cochrane BB, Lewis FM and Griffith KA (2011) Exploring a diffusion of benefit: Does a woman with breast cancer derive benefit from an intervention delivered to her partner? *Oncology Nursing Forum* 38(2): 207–214.

Derogatis L (2000) *BSI-18: Administration, scoring and procedures manual*. Baltimore, MD: Clinical Psychometric Research.

Donarelli Z, Gullo S, Lo Coco G, et al. (2015) Assessing infertility-related stress: The factor structure of the Fertility Problem Inventory in Italian couples undergoing infertility treatment. *Journal of Psychosomatic Obstetrics & Gynecology* 36: 58–65.

Donarelli Z, Coco G Lo, Gullo S, et al. (2016) The Fertility Quality of Life Questionnaire (FertiQoL) Relational subscale : psychometric properties and discriminant validity across gender. *Human Reproduction* 0(0): 1–11.

Eisenberg ML, Smith JF, Millstein SG, et al. (2010) Predictors of not pursuing infertility treatment after an infertility diagnosis: Examination of a prospective U.S. cohort. *Fertility and Sterility* 94: 2369–2371.

Emery M, Béran MD, Darwiche J, et al. (2003) Results from a prospective, randomized, controlled study evaluating the acceptability and effects of routine pre-IVF counselling. *Human Reproduction* 18: 2647–2653.

ESHRE Psychology and Counseling Guideline Development Group (2015) *Routine psychosocial care in infertility and medically assisted reproduction – A guide for fertility staff*. Available from: http://www.eshre.eu/~/media/Files/Guidelines/Psychology/ESHRE psychology guideline_2015_final_version 1_2.pdf.

Fertility Quality of Life tool (2016). Available from: http://sites.cardiff.ac.uk/fertiqol/download/ (accessed 15 April 2016).

Frasure-Smith N, Lespérance F and Talajic M (1995) The impact of negative emotions on prognosis following myocardial infarction: is it more than depression? *Health psychology : official journal of the Division of Health Psychology, American Psychological Association* 14(5): 388–398.

Gana K and Jakubowska S (2016) Relationship between infertility-related stress and emotional distress and marital satisfaction. *Journal of Health Psychology* 21(6): 1043–1054.

Gourounti K, Anagnostopoulos F and Vaslamatzis G (2010) Psychosocial Predictors of Infertility Related Stress: A Review. *Current Women's Health Reviews* 6: 318–331.

Greil a. L, McQuillan J and Sanchez D (2016) Does fertility-specific distress vary by race/ethnicity among a probability sample of women in the United States? *Journal of Health Psychology* 21(2): 183–192.

Greil AL (1997) Infertility and psychological distress: A critical review of the literature. *Social Science and Medicine* 45: 1679–1704.

HaCohen N, Amir D and Wiseman H (2016) Womens narratives of crisis and change: Transitioning from infertility to pregnancy. *Journal of Health Psychology*.

Hanna E and Gough B (2016) Searching for help online: An analysis of peer-to-peer posts on a male-only infertility forum. *Journal of health psychology*.

Hooper D, Coughlan J and Mullen M (2008) Structural equation modeling: guidelines for determining model fit. *Electronic Journal of Business Research Methods* 6: 53–60.

Hosmer D and Lemeshow S (2000) *Applied Logistic Regression*. 2nd edn. New York, USA: Jonh Wiley & Sons.

Hsu PY, Lin MW, Hwang JL, et al. (2013) The fertility quality of life (FertiQoL) questionnaire in Taiwanese infertilecouples. *Taiwanese Journal of Obstetrics and Gynecology*, Elsevier 52: 204–209.

Kahn JH (2006) Factor analysis in counseling psychology research, training, practice: Principles, advances, and applications. *The Counseling Psychologist* 34(5): 684–718.

Kenny D (1979) *Correlation and Causality*. New York, USA: Wiley.

Keramat A, Masoumi SZ, Mousavi S, et al. (2014) Quality of Life and its related factors in infertile couples. *Journal of Research in Health Sciences* 14: 59–66.

Kessler RC, Andrews G, Colpe LJ, et al. (2002) Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine* 32(6): 959–976.

Klonoff-Cohen H and Natarajan L (2004) The concerns during assisted reproductive technologies (CART) scale and pregnancy outcomes. *Fertility and Sterility* 81: 982–988.

Klonoff-Cohen H, Chu E, Natarajan L, et al. (2001) A prospective study of stress among women undergoing in vitro fertilization or gamete intrafallopian transfer. *Fertility and sterility* 76: 675–687.

Kvien TK, Heiberg T and Hagen KB (2007) Minimal clinically important improvement/difference (MCII/MCID) and patient acceptable symptom state (PASS): What do these concepts mean? *Annals of the rheumatic diseases* 66: 40–41.

Lancastle D and Boivin J (2005) Dispositional optimism, trait anxiety, and coping: unique or shared effects on biological response to fertility treatment? *Health psychology : official journal of the Division of Health Psychology, American Psychological Association* 24: 171–178.

Little TD, Cunningham W a., Shahar G, et al. (2002) To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling: A Multidisciplinary Journal* 9(2): 151–173.

Lopes V, Canavarro MC, Verhaak CM, et al. (2014) Are patients at risk for psychological maladjustment during fertility treatment less willing to comply with treatment? Results from the Portuguese validation of the SCREENIVF. *Human Reproduction* 29: 293–302.

Luk BH-K and Loke AY (2015) The Impact of Infertility on the Psychological Well-Being, Marital Relationships, Sexual Relationships, and Quality of Life of Couples: A Systematic Review. *Journal of Sex & Marital Therapy* 41(6): 610–625.

Martins M, Peterson B, Almeida V, et al. (2011) Direct and indirect effects of perceived social support on womens infertility-related stress. *Human Reproduction* 26: 2113–2121.

Martins M, Peterson B, Costa P, et al. (2013) Interactive effects of social support and disclosure on fertility-related stress. *Journal of Social and Personal Relationships* 30(4): 371–388.

Martins M, Basto-Pereira M, Pedro J, et al. (2016) Male psychological adaptation to unsuccessful medically assisted reproduction treatments: a systematic review. *Human Reproduction Update* 22: 466–478.

Melo C, Gameiro S, Canavarro MC, et al. (2012) Does the FertiQol assess quality of life? Results from the validation of the Portuguese version of the FertiQol. *Human Reproduction* 27(suppl 2): ii268–ii273.

Moura-Ramos M, Gameiro S, Canavarro MC, et al. (2012) Assessing infertility stress: Re-examining the factor structure of the Fertility Problem Inventory. *Human Reproduction* 27: 496–502.

Newton CR, Sherrard W and Glavac I (1999) The fertility problem inventory: Measuring perceived infertility- related stress. *Fertility and Sterility* 72: 54–62.

Pallant J (2010) *SPSS Survival Manual*. 4th edn. McGraw-Hill (ed.), Maidenhead.

Peterson B, Newton CR and Rosen KH (2003) Examining congruence between partners´perceived infertility-related stress and its relationship to marital adjustment and depression in infertile couples. *Family Process* 42: 59–70.

Peterson B, Newton C, Rosen K, et al. (2006) Gender differences in how men and women who are referred for IVF cope with infertility stress. *Human Reproduction* 21: 2443–2449.

Peterson B, Newton CR, Rosen KH, et al. (2006) The relationship between coping and depression in men and women referred for in vitro fertilization. *Fertility and Sterility* 85: 802–804.

Salvatore P, Gariboldi S, Offidani A, et al. (2001) Psychopathology, personality, and marital relationship in patients undergoing in vitro fertilization procedures. *Fertility and Sterility* 75: 1119–1125.

Schmidt L (2006) Infertility and assisted reproduction in Denmark. Epidemiology and psychosocial consequences. *Danish medical bulletin* 53: 390–417.

Schmidt L (2009) Social and psychological consequences of infertility and assisted reproduction - what are the research priorities? *Human Fertility* 12: 14–20.

Schmidt L, Holstein BE, Boivin J, et al. (2003) Patients' attitudes to medical and psychosocial aspects of care in fertility clinics: Findings from the Copenhagen Multi-centre Psychosocial Infertility (COMPI) Research Programme. *Human Reproduction* 18: 628–637.

Schmidt L, Holstein BE, Christensen U, et al. (2005) Communication and coping as predictors of fertility problem stress: Cohort study of 816 participants who did not achieve a delivery after 12 months of fertility treatment. *Human Reproduction* 20: 3248–3256.

Schmidt L, Christensen U and Holstein BE (2005) The social epidemiology of coping with infertility. *Human Reproduction* 20: 1044–1052.

Slade P, Emery J and Lieberman BA (1997) A prospective, longitudinal study of emotions and relationships in in-vitro fertilization treatment. *Human reproduction (Oxford, England)* 12: 183–190.

Slade P, O'Neill C, Simpson A., et al. (2007) The relationship between perceived stigma, disclosure patterns, support and distress in new attendees at an infertility clinic. *Human Reproduction* 22: 2309–2317.

Smeenk JM, Verhaak CM, Eugster A, et al. (2001) The effect of anxiety and depression on the outcome of in-vitro fertilization. *Human reproduction* 16: 1420–1423.

Spielberger C, Catalano R and Dooley D (1983) *Manual for the State–Trait Anxiety Inventory STAI (Form Y)*. Palo Alto, CA: Consulting Psychologists Press.

Statistics Denmark (2015) Vital Statistics. Available from: https://www.dst.dk/en (accessed 6 October 2016).

Streiner DL and Norman GR (1995) *Health Measurement scales*. 2nd ed. Oxford, UK: Oxford University Press.

Streiner DL and Norman GR (2008) *Health measurement scales: a practical guide to their development and use.* Oxford, UK: Oxford University Press.

Verberg MFG, Eijkemans MJC, Heijnen EMEW, et al. (2008) Why do couples drop-out from IVF treatment? A prospective cohort study. *Human Reproduction* 23: 2050–2055.

Verhaak C, Lintsen AME, Evers AWM, et al. (2010) Who is at risk of emotional problems and how do you know? Screening of women going for IVF treatment. *Human Reproduction* 25: 1234–1240.

Youden WJ (1950) Index for rating diagnostic tests. *Cancer* 3(1): 32–35.

Zachariae R, Melchiorsen H, Frøbert O, et al. (2001) Experimental pain and psychologic status of patients with chest pain with normal coronary arteries or ischemic heart disease. *American heart journal* 142: 63–71.

Zachariae R, Zachariae C, Ibsen HHW, et al. (2004) Psychological symptoms and quality of life of dermatology outpatients and hospitalized dermatology patients. *Acta Dermato-Venereologica* 84: 205–212.

Table 1. Correlations between Total Scores and Respective Subscales, means and standardized values of the Modified COMPI-FPSS, FPI, FertiQoL, and BDI-II and STAI-State.

	FPSS Soc	FPSS Mar	FPSS Per	FPSS total	FPI Soc	FPI Sex	FPI Rel	FPI Rcl	FPI Nfp	FPI total	FQoL Emo	FQoL M/B	FQoL Rel	FQoL Soc	FQoL Env	FQoL Tol	FQoL total	BDI total	STAI State
FPSS Soc	-																		
FPSS Mar	.44**	-																	
FPSS Per	.67**	.44**	-																
FPSS total	.83**	.72**	.91**	-															
FPI Soc	.66**	.38**	.80**	.77**	-														
FPI Sex	.44**	.48**	.53**	.59**	.55**	-													
FPI Rel	.32**	.57**	.38**	.50**	.41**	.44**	-												
FPI Rcl	.32**	.24**	.46**	.43**	.42**	.30**	.19**	-											
FPI Nfp	.36**	.26**	.55**	.50**	.54**	.35**	.27**	.72**	-										
FPI total	.57**	.50**	.75**	.76**	.82**	.67**	.60**	.73**	.83**	-									
FQoL Emo	-.58**	-.44**	-.85**	-.80**	-.73**	-.51**	-.43**	-.50**	-.59**	-.76**	-								
FQoL M/B	-.52**	-.42**	-.77**	-.73**	-.61**	-.53**	-.39**	-.44**	-.50**	-.67**	.75**	-							
FQoL Rel	-.27**	-.66**	-.29**	-.47**	-.31**	-.52**	-.62**	-.12	-.11	-.42**	.31**	.34**	-						

	FPSS Soc	FPSS Mar	FPSS Per	FPSS total	FPI Soc	FPI Sex	FPI Rel	FPI Rcl	FPI Nfp	FPI total	FQoL Emo	FQoL M/B	FQoL Rel	FQoL Soc	FQoL Env	FQoL Tol	FQoL total	BDI total	STAI State
FQoL Soc	-.63**	-.40**	-.77**	-.76**	-.73**	-.52**	-.41**	-.48**	-.49**	-.72**	.78**	.77**	.33**	-					
FQoL Env	-.15*	-.27**	-.07	-.18**	-.08**	-.09	-.36**	-.07	-.03	-.16**	.17**	.16**	.33**	.17**	-				
FQoL Tol	-.42**	-.40**	-.64**	-.62**	-.50**	-.36**	-.35**	-.29**	-.35**	-.51**	.67*	.66**	.27**	.59**	.14*	-			
FQoL total	-.60**	-.57**	-.82**	-.83**	-.70**	-.58**	-.56**	.47**	-.51**	-.76**	.89**	.88**	.54**	.85**	.42**	.73**	-		
BDI total	.48**	.38**	.70**	.66**	.57**	.56**	.42**	.42**	.46**	.65**	-.72**	-.77**	-.32**	-.70**	-.09	-.56**	-.76**	-	
STAI State	.43**	.35**	.64**	.60**	.52**	.50**	.42**	.34**	.41**	.59**	-.69**	-.73**	-.30**	-.65**	-.14*	-.54**	-.73**	.86**	-
Mean	2.65	4.66	9.09	16.40	24.17	11.59	18.47	23.35	40.29	117.87	14.99	14.61	12.41	9.48	5.27	11.34	71.10	10.05	68.00
(SD)	(2.78)	(3.09)	(4.99)	(2.78)	(10.29)	(5.94)	(7.66)	(8.00)	(10.49)	(31.49)	(5.39)	(5.04)	(2.56)	(2.47)	(1.69)	(3.09)	(15.25)	(8.55)	(10.92)

Note. ** P<0.001; * P<0.05. FPSS Soc, Social domain; FPSS Mar, Marital domain; FPSS Per, Personal domain; FPSS total, Fertility problem stress scales total score. FPI Soc, Social concern; FPI Sex, Sexual concern; FPI Rel, Relational concern; FPI Rcl, Rejection of a childfree-lifestyle; FPI Nfp, Need for parenthood; FPI total, Fertility Problem inventory total core. FQoL Emo, Emotional subscale; FQoL M/B, Mind-body subscale; FQoL Rel, Relational subscale; FQoL Soc, Social subscale; FQoL Env, Environment subscale; FQoL Tol, Tolerability subscale; FQoL total, Fertility quality of life total score. BDI total, Beck Depression Inventory total score. STAI State, State anxiety subscale.