Is infertility a risk factor for impaired male fertility?

M. Pook¹, B. Tuschen-Caffier² and W. Krause³

¹Department of Psychology, University of Siegen, ²Department of Psychology, University of Bielefeld and ³Clinic of Andrology, Philipps-University of Marburg, Germany

BACKGROUND: Previous research has suggested an interaction between distress and male fertility. The present longitudinal study sought to deliver evidence for a negative impact of distress due to infertility on sperm concentration.

METHODS: The sample consisted of 120 patients who twice visited an andrological clinic on their own initiative for fertility work-ups. Baseline and follow-up examinations were at least 6 months apart. Prior to each fertility work-up, patients completed a questionnaire assessing distress due to infertility.

RESULTS: Path analyses revealed that the level of infertility distress at follow-up has a negative impact on the change in sperm quality from baseline to follow-up assessment. Distress scores were highly stable. As a consequence, the level of distress at baseline assessment provided only little additional information for the changes in sperm concentration. Further analysis suggested that the fertility status had no impact on infertility distress. CONCLUSION: The present study delivers the strongest evidence to date that distress due to infertility is a significant risk factor for a decrease in sperm quality.

Key words: counselling/male infertility/psychological stress/sperm count

Introduction

When aetiologies for testicular dysfunction are discussed, psychological stress is often also considered to be relevant. While initial reports about the impact of distress on male fertility were rather anecdotal, in the last few decades more systematic research has been conducted. The correlation between sperm quality and distress has been computed in several studies. Given the wide range of values of sperm parameters, it is not too surprising that these cross-sectional studies led to contradictory results (see Pook et al., 1999a). Quite obviously, categories such as cryptozoospermia (<2 × 10⁶ spermatozoa per ml), oligozoospermia (2–20 × 10⁶/ml), normozoospermia (20–250 × 10⁶/ml) or polyzoospermia (>250 × 10⁶/ml) do not simply reflect a specific stress level.

Other research has focused on the change of sperm quality in men exposed to stress, rather than on a single semen specimen of participants. For example, it has been found that sperm quality decreased subsequent to a natural disaster such as an earthquake (Fukuda et al., 1996). Some well-controlled studies have also revealed that basic sperm parameters declined while undergoing IVF treatment (Harrison et al., 1987; Ragni and Caccamo, 1992; Clarke et al., 1999). These studies were based on the rationale that medical examinations and involvement in the treatment of the spouse are stressful for the majority of patients. Therefore, a decline in sperm quality was hypothesized when comparing the last pre-treatment semen analysis with the specimen from the treatment cycle several weeks later.

In the studies on IVF patients, there is another noteworthy aspect. An interaction between stress and fertility is more or less explicitly assumed: sperm quality is an essential factor for male fertility, while the condition of infertility (or its treatment) causes a decline in sperm quality. The studies on IVF patients, however, lack an appropriate psychological assessment. Changes of sperm quality were set in relation to the patients’ answers on psychometric tests in only one study (Clarke et al., 1999). In that study, patients rated aspects related to giving a semen specimen (such as stress or anxiety) immediately after having delivered the specimen. Short lasting stressors such as giving a semen specimen, however, cannot be expected to cause changes in sperm quality ( McGrady, 1984). Moreover, even the assumption that undergoing IVF treatment per se causes distress is not very conclusive. The fact that stressful situations are not necessarily associated strongly with perceived stress or negative affect (Cohen et al., 1993) should be taken into account here.

Therefore, the design used in the research on IVF patients was extended in a study on infertile males returning for a new andrological examination, at least 6 months after their previous fertility work-up (Pook et al., 1999a). Prior to the second work-up, participants completed a questionnaire on chronic infertility distress. The scores had predictive value for changes in sperm quality from the previous specimen to the one delivered subsequent to completing the questionnaire. However, psychological assessment was still imperfect because there were no baseline distress scores. Therefore, more evidence is needed...
revealed and confirmed that there is a single dimension underlying the extensively (for a review see Pook and Krause, 2002). Factor analyses of the desire for a child. The Infertility Distress Scale has been evaluated infertility; (vii) the frequency of thoughts about infertility; and (viii) that infertility represents a threat; (vi) feelings of helplessness due to infertility as whole; (iii) the importance of a child; (iv) the of eight items on 5-point Likert scales. The items ask for ratings of (i) resulting from infertility as perceived by the participants. It consisted of a total of 120 patients were included in the study sample. Mean age of the patients at the initial fertility work-up was 32.8 years (range 20.6–42.3). Mean duration of infertility at the initial fertility work-up was 30.9 months (range 2.0–120.0). Mean interval between the initial and the subsequent fertility work-up was 21.6 months (range 6.0–73.5).

Despite the exclusion criteria outlined above, participants presumably represent typical attendees at the andrological clinic. That means that ~50% were self-referred, while the other half are referred by general practitioners or other medical services. The initial fertility work-up was the patients’ first at the andrological clinic where the present study was conducted. While 50 patients (41.7%) underwent some fertility diagnostics before visiting the clinic, the initial work-up was the first complete fertility diagnostics for the majority of participants. After the initial work-up, 61 patients (50.8%) received the diagnosis of male infertility.

Variables

The Infertility Distress Scale was employed to assess the stress resulting from infertility as perceived by the participants. It consisted of eight items on 5-point Likert scales. The items ask for ratings of (i) the distress due to the spouse’s last menstruation; (ii) the distress due to infertility as a whole; (iii) the importance of a child; (iv) the perception that infertility represents a challenge; (v) the perception that infertility represents a threat; (vi) feelings of helplessness due to infertility; (vii) the frequency of thoughts about infertility; and (viii) the desire for a child. The Infertility Distress Scale has been evaluated extensively (for a review see Pook and Krause, 2002). Factor analyses revealed and confirmed that there is a single dimension underlying the eight items. Internal consistency (α = 0.89) and re-test reliability (e.g. \( r = 0.74 \) for a follow-up of 4 months) were found to be good. The scale correlated highly with an already validated measure of infertility distress. Correlations with widely used measures of depressiveness are in the small to medium range. Additional findings indicate that the Infertility Distress Scale is sensitive to change. There are also norms available for the Infertility Distress Scale based on a sample of >750 patients. Even in the large normative sample, the scores were distributed approximately normally.

Sperm concentration was chosen as an indicator for sperm quality, because previous research (Harrison et al., 1987; Ragni and Caccamo, 1992; Clarke et al., 1999; Pook et al., 1999a) has suggested that this parameter in particular is sensitive to distress. Since the distribution of sperm concentration is skewed in general, cube root transformation was performed to achieve normality (Mallidis et al., 1991). For estimation of sperm concentration, guidelines of the World Health Organization (1999) were followed. Since pre-ejaculation abstinence can have a serious impact on sperm concentration (Poland et al., 1985), this variable was also included in the analysis.

Procedure

Over a period of 3 years, prospective data of consecutive repeat consulters were collected from an andrology clinic. When designing the study, 3 years were estimated to be sufficient for gathering data of 100–150 subjects, which is assumed to be the minimum sample size for the planned (and performed) data analysis. All appointments for fertility work-up were made on the participants’ initiative. Upon arriving at the andrology clinic, the patients filled out the Infertility Distress Scale prior to the medical examination as a matter of routine. The medical examination was the same for every patient. It consisted of a medical history, a physical investigation including sonography of the testes, and a semen analysis. After all laboratory tests had been performed, the patients received the reports of the fertility work-up in written form by mail. In accordance with the local ethics committee, patients’ consent was not necessary for the present study because the (anonymized) data were obtained during routine clinical investigation.

Data analysis

Different models of the relationship between distress and sperm quality were tested and compared using path analysis. Path analysis estimates the relative importance of a path of influence and, therefore, provides information on how well a hypothesized causal model fits empirical data. In general, path diagrams are provided for visual representation of the cause–effect relationships amongst variables. In these diagrams, an arrow (= path) indicates a causal relationship between two variables, while the path coefficient is the numerical estimate of the causal relationship. Path coefficients are similar to correlation coefficients: both indicate the strength of an association, while a negative sign specifies an inverse relationship between two variables. However, in contrast to correlation coefficients, path coefficients might, under specific circumstances, exceed the value of 1.

Statistical significance of path coefficients indicates the usefulness of a hypothesized causal model. Causal models can also be evaluated using the \( \chi^2 \) statistic, the goodness of fit index (GFI), adjusted goodness of fit index (AGFI) and the root mean square residual (RMSR). The \( \chi^2 \) statistic tests the hypothesis that the path model is statistically different from the saturated model, where all possible relationships are specified. The GFI and AGFI provide an indication of how well the hypothesized model fits the data in comparison to a null model, which assumes no relationships among the variables. The RMSR indicates the degree of goodness of fit of the hypothesized model in comparison to a baseline model.

Materials and methods

Sample

The sample consisted of patients who visited an andrological clinic for two complete fertility work-ups at least 6 months apart. Exclusion criteria were (i) fathering a child between the initial and the subsequent fertility work-up; (ii) an acute testicular dysfunction caused by an accident, an operation, a cytotoxic therapy or similar causes; (iii) azoospermia; and (iv) participation in a previous study (Pook et al., 1999a). A total of 120 patients were included in the study sample. Mean age of the patients at the initial fertility work-up was 32.8 years (range 20.6–42.3). Mean duration of infertility at the initial fertility work-up was 30.9 months (range 2.0–120.0). Mean interval between the initial and the subsequent fertility work-up was 21.6 months (range 6.0–73.5).

Descriptive statistics

Table I. Descriptive statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Baseline assessment</th>
<th>Follow-up assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Sperm concentration (×10⁹/ml)</td>
<td>56.85</td>
<td>72.00</td>
</tr>
<tr>
<td>Motility (percentage of motile sperm)</td>
<td>35.18</td>
<td>21.91</td>
</tr>
<tr>
<td>Morphology (percentage of deformed sperm)</td>
<td>83.56</td>
<td>9.25</td>
</tr>
<tr>
<td>Sperm concentration (Cube root transformed)</td>
<td>3.23</td>
<td>1.52</td>
</tr>
<tr>
<td>Infertility distress</td>
<td>16.58</td>
<td>6.76</td>
</tr>
</tbody>
</table>

*aVariables were normally distributed and included in the path analysis.

Infertility distress and sperm quality

Table I. Descriptive statistics

![Table I. Descriptive statistics](https://academic.oup.com/humrep/article/19/4/954/2913639)

955
consistent with the pattern of covariation among the observed variables; statistical significance of a $\chi^2$ statistic indicates poor fit. GFI and AGFI test the extent to which the path model was consistent with the data; here, values >0.90 indicate good fit. The RMSR evaluates the closeness of the original covariance matrix to the reproduced covariance matrix; values <0.05 indicate appropriate closeness. In addition, $D\chi^2$ can be used for direct comparisons of competing path models; statistical significance of $D\chi^2$ indicates the superiority of one model over the other.

Detailed information on path analysis can be found in books on structural equation modelling (e.g. Byrne, 1998). LISREL (version 8.30) and SPSS for Windows (release 10.0.5) were applied for data analysis in the present study.

Results

Descriptive statistics are given in Table I. Different models concerning the impact of distress on sperm quality were evaluated. The (simplified) structural models, the path coefficients and the fit indices of these models are presented in Figure 1.

Model A

The most conservative model proposes that both infertility distress and sperm quality show some stability over time. This model does not consider any impact of distress on sperm quality. In fact, the path coefficients indicated that both infertility distress and sperm quality were highly stable. Accordingly, fit indices suggested that the empirical data reasonably fit Model A.

Model B

An extended model proposes that infertility distress and sperm quality show stability over time, while the level of stress at follow-up has impact on sperm quality at follow-up. The assumed impact of distress was supported by the significant path coefficient ($T = -2.80$) from distress at follow-up to sperm quality at follow-up. Moreover, excellent fit indices, along with a significantly better performance [$D\chi^2(1) = 7.35, p < 0.01$], indicated that Model B is more appropriate than Model A.

Model C

High stability of infertility distress could lead to the idea that the questionnaire used merely assesses a personality trait and therefore the score is not associated too closely with a patient’s current situation. If this were the case, distress at baseline should be relevant for sperm quality at follow-up, even if distress at follow-up is regarded as irrelevant. In Model C, the path from distress at baseline to sperm quality at follow-up was not significant ($T = -0.06$). More importantly, fit of Model C
was significantly poorer than fit of Model B \[\Delta \chi^2(1) = 6.15, P < 0.05\]. Moreover, Model C did not even perform significantly better that Model A \[\Delta \chi^2(1) = 1.2, \text{NS}\]. Thus, the timing of assessing infertility distress is more relevant than potential trait-like aspects of the scale’s content.

**Model D**

This model proposes that sperm quality and infertility distress show stability over time, and that the distress at baseline has an impact on sperm quality at follow-up, in addition to the distress at follow-up. If the signs of the latter two path coefficients are opposite, the model would indicate that the change in distress is highly relevant for the sperm quality at follow-up. However, the path from distress at baseline to sperm quality at follow-up was not significant \(T = 1.23\), while the path from distress at follow-up was \(T = -2.84\). Fit indices were nearly optimal. Model D performed significantly better than Model A \[\Delta \chi^2(2) = 8.88, P < 0.05\], but not better than Model B \[\Delta \chi^2(1) = 1.53, \text{NS}\].

To summarize the findings up to now, Model B and Model D—each assuming that the level of infertility distress at follow-up has some impact on the change in sperm quality—performed best. Due to the already excellent fit of Model B, it is not possible for Model D to fit significantly better than Model B. Since the insignificant fit improvement, as well as the insignificant path coefficient from distress at baseline to sperm quality at follow-up in Model D, might simply be a problem of sample size, the unique contribution (‘explained variance’) of each variable was estimated. While sperm quality at baseline explained 64.1\% of the variance of sperm quality at follow-up, distress at follow-up explained 2.3\% and distress at baseline 0.4\% additionally. Therefore, Model B seems to be sufficient for explaining the impact of infertility distress on sperm quality.

Next, associations according to Model B were controlled for pre-ejaculation abstinence. Inclusion of abstinence (Model E) led to no significant change in model fit \[\Delta \chi^2(5) = 5.84, \text{NS}\]. In line with the known impact of abstinence, path coefficients indicated that increasing abstinence is associated with increased sperm quality. However, it cannot be assumed that the impact of distress on sperm quality is an artefact due to pre-ejaculation abstinence, because the path from distress at follow-up to sperm quality at follow-up was still significant \(T = -2.36\).

In order to explore further the impact of distress, changes in sperm concentration were contrasted descriptively for patients either seriously or non-seriously distressed at follow-up. According to previous research, patients with a score of \(\geq 21\) on the Infertility Distress Scale have to be regarded as seriously distressed (Pook and Krause, 2002). While the mean sperm concentration was stable in the group of non-seriously distressed patients, there was a decline of 33.3\% in the group of highly distressed patients (Table II). At follow-up, however, both of the means were still in the range of normozoospermia.

Beyond the main analysis of the present study, the aspect of an interaction between distress and male fertility was then explored further. The findings up to now suggest that distress due to infertility has a negative impact on an essential parameter of male fertility. Now we examine whether the information about this parameter also influences the level of infertility distress. Therefore, in Model F, Model B was extended by a path going from sperm quality at baseline to distress at follow-up. The respective path coefficient was insignificant \(T = -0.31\) and the fit improvement was marginal \[\Delta \chi^2(1) = 0.10, \text{NS}\]. Thus, the assumption that knowledge about a patient’s sperm concentration affects his infertility distress is not supported.

Finally, we analysed whether receiving the diagnosis of male infertility on the basis of the complete fertility work-up at baseline induces greater distress. To do so, a two-way analysis of variance was conducted, with fertility status as a between-subject factor (male infertility versus unimpaired fertility) and time (baseline versus follow-up assessment) as a within-subject factor. The dependent variable was infertility distress. The central point of interest was the interaction between fertility status and time. Since the interaction was insignificant \([F(1,118) < 1; \eta^2 < 0.001]\), the diagnosis of male infertility does not seem to be an additional source of distress for male fertility patients.

**Discussion**

At a time when an ongoing controversy about a global decrease in sperm concentration (Carlsen et al., 1992) seems to dominate the discussion, the present study follows a more individual and psychological focus. It explores whether distress due to infertility has a negative impact on an essential parameter of male fertility. Indeed, the path analyses confirmed the assumed association. Thus, when considering the content of the employed measure of infertility distress, the conclusion one could reach is paradoxical: the more a patient desires a child, the more relevant a child is for him and the more often he thinks about his infertility, the higher the risk for impaired sperm quality. However, before arriving at such a conclusion, the present findings need to be discussed carefully.
It is the first time that a longitudinal study on the impact of any kind of stress on sperm quality has been conducted with a psychometrically sound assessment prior to both initial and follow-up semen analysis. Nevertheless, in the present study, the baseline scores delivered no additional information of relevance. Two aspects should be considered here for post hoc explanation. First, the length of follow-up is not a fixed period, and this variation could have limited the impact of baseline scores. Secondly, the surprisingly high stability of distress scores also restricts the impact of the initial scores. However, any assumption that infertility distress merely reflects a personality trait is unwarranted, since the respective model did not perform well. This result is supported by previous findings that the employed Infertility Distress Scale is sensitive to change (Ochsendorf et al., 2002; Pook et al., 2002).

The path analyses suggest that the current level of infertility distress has a significant impact on the change in sperm quality. One might think that the direction of the arrow between distress and sperm quality at follow-up is arbitrary, because the path is based on a cross-sectional correlation. However, one should be aware that all participants provided the semen specimen after completing the questionnaire. Thus, when filling out the questionnaire, the patients were blind to their current fertility status. As a consequence, it is not reasonable to assume any other association between distress and sperm quality at follow-up that is different from the relationship in Model B.

Model B is in line with the findings of a previous study (Pook et al., 1999a) in which, however, less conclusive statistics were employed, a smaller sample was studied and less information about the patients was available. In the previous study, the need for a replication study was noted, because artefacts due to uncontrolled variables (e.g. starting or giving-up cigarette smoking) could not be ruled out. Since the present and the previous study led to similar results, it became implausible that any chance impact might account for the association between distress and the changes in sperm quality. The present analysis goes far beyond the previous study in quantifying the unique contribution of distress to the variance of sperm quality at the follow-up assessment. In general, if a variable explains a mere 2.3% variance of another variable, its impact has to be considered small. So, even if the present analysis suggests that the impact of infertility distress could be as relevant as the widely accepted impact of pre-ejaculation abstinence (Poland et al., 1985), the question arises as to whether there is reason to be concerned about the consequences of infertility distress on male fertility.

Yet even a very small amount of explained variance can be highly relevant. For example, mathematical models presented in the debate about global changes in sperm quality and supporting either a recent increase or a decrease sometimes even differ in the range of <2% of explained variance (Olsen et al., 1995). In addition, one must be aware that the proportion of explained variance does not say very much about the proportion of patients whose fertility is impaired by infertility distress. Taking its chronicity into account, infertility distress must have already influenced the baseline assessment of sperm quality in the present study. Only a prospective study, in which the baseline assessment of sperm quality takes place before the participants actively try to father a child, could identify the exact proportion of patients whose fertility is impaired due to infertility distress. Such a study, however, would probably be prohibitive because of costs. Nevertheless, it is obvious that it would be rather misleading to focus simply on the small proportion of explained variance in the present analysis.

Considering the high stability of distress scores, findings on the course of distress after fertility work-up are also relevant. Previous research has revealed that after finishing fertility diagnostics, male patients felt less distressed (Connolly et al., 1992). In a more detailed study, however, it was found that distress declined significantly only in first-time consultants, but not in repeat consultants (Pook et al., 2002). Since the participants in the present study were repeat consultants at the follow-up assessment, it cannot be assumed that they felt much less distressed after finishing the fertility work-up. Therefore, they are at risk for a further decline in sperm quality. Moreover, iatrogenic distress caused by future interventions must be considered. As pointed out above, undergoing IVF treatment obviously induces distress. Unfortunately, treatment failure also represents a serious stressor. Longitudinal studies indicated that distress rises sharply 3 weeks (Newton et al., 1990) and even 6 months (Slade et al., 1997) after unsuccessful IVF treatment. Even psychological interventions aiming at a reduction of distress can lead to rather unintended effects. In a randomized controlled study (Takefman et al., 1990), two groups of patients received extensive information about problems related to infertility and about coping possibilities. Those patients showed poorer adjustment compared with patients who had not received the extensive information. For the purpose of judging these findings, it is interesting to note that extended efforts to cope with infertility have been found to be associated with both high distress and a subsequent decline in sperm quality (Pook et al., 1999b).

While there are different sources that cause iatrogenic distress in fertility patients, the present findings indicate that the diagnosis of impaired fertility does not represent such a source. To the best of our knowledge, there are two prospective studies available that focus specifically on the impact of fertility diagnosis in male patients. One study (Connolly et al., 1992) led to rather inconclusive results. On the one hand, the data suggested that males receiving the diagnosis of impaired fertility are more distressed at follow-up. On the other hand, there was no significant interaction between diagnosis and time, so the study could not deliver evidence for differences between groups with different diagnoses with respect to the course of distress. In a more recent study (Pook et al., 2002), no impact of the fertility status on infertility distress was detected. Thus, the present findings do not contradict previous research. For us, the absence of increased distress in the male infertility subgroup is not implausible. Of course, it can be distressing for some patients to receive the diagnosis of male infertility. However, receiving the diagnosis of male infertility may be no more distressing than achieving no pregnancy with ‘normal’ fertility.

Some limitations of the present study should be noted. First and foremost, only one of the numerous sperm
variables has been analysed. Unfortunately, given the restricted sample size, any additional variable would have endangered the accuracy and stability of path models as they depend on the ratio of subjects and estimated parameters. At present, only in Model E is the ratio below 1:10, which means this model should be used with some caution. If additional sperm variables were included in the path models, the data analysis could not deliver any reliable information at all. Therefore, analysing only one basic sperm parameter appeared to be the only option for studying the association between stress and male fertility with a limited patient sample. The sample size also prevented the formation of subgroups (e.g. long versus short time between clinic visits) and a subsequent comparison of path models for such subgroups. As a consequence, the current paths models have to be considered valid for all repeat consulters of andrological services that are similar to the present sample.

Our study is limited additionally by the uncertainty about the mechanisms mediating the impact of stress on spermatogenesis. Due to insufficient information, it cannot even be ruled out for the present sample that increased distress led to increased cigarette smoking, which then negatively affected spermatogenesis. However, during the extensive evaluation of the employed Infertility Distress Scale (Pook and Krause, 2002), no association between the score and aspects such as nicotine consumption, age, duration of infertility, etc. has been detected. Thus, it seems unlikely that smoking is the central mediator between infertility distress and sperm quality.

To summarize, the present study was not designed to explore the proportion of patients whose fertility is impaired due to infertility distress. Nevertheless, it has delivered the strongest evidence to date that distress due to infertility is a significant risk factor for a decrease in sperm quality. At present, the mechanisms mediating the impact of stress on spermatogenesis are unclear. Additional research is needed here. However, it seems to be more urgent to study how the well-being of infertile males can be improved. To the best of our knowledge, there have been no randomized controlled studies in which a beneficial effect of psychological services for fertility patients could be shown. Moreover, in general, even the uptake rates of such services are low (Boivin, 1997). Therefore, evaluation of carefully designed psychological interventions must go hand-in-hand with improved recruitment strategies (Pook et al., 2001). It could also be worthwhile to evaluate the clinical routine for fertility services, as well as procedural alternatives with respect to the impact on the patient’s well-being. We hope that the present study might contribute to initiate research into these matters.

References


Submitted on November 24, 2003; accepted on January 8, 2004