Introduction

Infertility is defined as the absence of conception after 12 months of sexual activity without the use of contraceptives (WHO, 1993). Its prevalence in the population varies among different countries, from a minimum of 5% to a maximum of 30% of all couples (WHO, 2002). Of these, approximately 10% do not resolve spontaneously, and many of them seek medical help, being eventually offered assisted reproduction techniques when infertility persists.

It is to be expected that any difficulty met by a couple in this field will impose a considerable psycho-social burden. Indeed, throughout history, infertility has been a problem of such magnitude that it has acquired symbolic and religious connotations (Mazure et al., 1992; Greil, 1997; Bahadur et al., 2001; Pennings, 2002; Baetens et al., 2003; Jones and McMahon, 2003; King, 2003). Knowledge of one’s own reproductive power is a basic element of self-esteem for both men and women (Pines, 1990).

The psychological implications related to human infertility have been extensively studied (Pines, 1990; Moller and Fallstrom, 1991; Marari et al., 1992; Greil, 1997; Ardanti et al., 1999; Oddens et al., 1999). The most frequent possible disturbances are anxiety, depression, low self-esteem, strong psychological tension, crisis within the marital relationship, separation, and divorce (Boivin et al., 1998; Demyttenaere et al., 1998; Oddens et al., 1999; Lee et al., 2001).

The strain of treatment is also a factor in psycho-emotional maladjustment. There is no doubt that assisted reproductive techniques have expanded the opportunities for the treatment of infertility. However, the sophisticated technology involved in assisted reproduction creates an atmosphere of anguish related to the costs and the relatively low success rates (Marari

Comparison of the psychological evaluation test and classical psychoanalysis in infertile women

Dr Franco Jr worked from 1984 to 1985 in the University of California, Irvine, USA. In 1988 he became the Scientific Director of the Centre for Human Reproduction of Maternity Sinhá Junqueira in Ribeirão Preto, Brazil. He was a member of the Brazilian Medicine Federal Council which has defined ethical rules for using assisted reproduction techniques. He was also one of the founders of the Assisted Reproduction Brazilian Society and the founder of the Jornal Brasileiro de Reprodução Assistida, where he was the responsible editor from 1997 to 2002. Among his published papers, 176 were published in Brazil and 60 in other countries. At present he is the Executive Director of the Latin America Assisted Reproduction Net. His special interests are stress and IVF, and ovarian stimulation.

Dr Franco Jr

JG Franco Jr1,2,3, PG Silva1, RS Achê1,2, RR Baruffi1, AL Mauri1, CG Petersen1, V Felipe1, L Vagnini1, JBA Oliveira1
1Centre for Human Reproduction, Sinhá Junqueira Maternity Foundation–Ribeirão Preto SP, Brazil; 2Department of Gynecology and Obstetrics, Faculty of Medicine of Botucatu (UNESP) SP, Brazil
3Correspondence: Centro de Reprodução Humana, Fundação Maternidade Sinhá Junqueira, Rua D. Alberto Gonçalves 1500, CEP 14085-100 Ribeirão Preto SP, Brazil. Tel: +55 16 6262909/628 1196/615 8230; Fax: +55 16 6283755; e-mail: crh@crh.com.br and/or franco@crh.com.br

Abstract

This study aims to compare a psychological evaluation test to classical psychoanalysis in infertile women. Two hundred women were submitted to the Psychological Evaluation Test (PET). The sum of the scores for the responses ranged from 15 to 60 points, with scores ≥30 points being defined as ‘psycho-emotional maladjustment’ (cut-off point: median + 25%). For comparison, the patients were simultaneously submitted to a psychological examination by a psychologist, who was unaware of the PET results. Of the 200 patients, 66 (33%) presented a test with ≥30 points (‘psycho-emotional maladjustment’) and 134 (67%) a test with <30 points (normal). Upon psychological examination, 105 (52.5%) presented an abnormal evaluation and 95 (47.5%) a normal evaluation. For the PET, statistical analysis showed 82% efficiency, 62% sensitivity, 98% positive predictive value, 99% specificity, 70% negative predictive value, likelihood ratio for a positive test result 62, and likelihood ratio for negative test result 0.38. The PET proved to be a useful clinical instrument, being of help in the selection of patients with psychological needs induced by infertility.

Keywords: infertility, psychological evaluation, psychological test, score
Among all types of treatment of infertility, IVF is the one involving the highest stress, with up to 80% of patients presenting a moderate or marked level of stress (Hynes et al., 1992; Connolly et al., 1993). Many couples also report acute depression after failure of one cycle (Marari et al. 1992; Covington, 1997). This depression affects women more than men, persisting for more than 6 months (Domar et al., 1992; Hunt and Monach, 1997; Slade et al., 1997; Guerra et al., 1998; Hammarberg et al., 2001).

It seems clear that, due to the emotional consequences of the medical treatment, patients will request psychological support. So, the clinician has a problem when managing infertile patients. The situation is exacerbated particularly when working in services which do not have their own psychology sector or scarcity of staff or time pressures prevent the proper evaluation of all patients seeking help.

The ability to adapt to stress or to potentially anguishing stimuli depends on the personality of the individual and on her defence mechanisms (Connolly et al., 1992; Golombok, 1992; Glover et al., 1999; Covington, 2001; Anderson et al., 2003). For individuals to adjust, they must go through cognitive, emotional and behavioural changes. If, on the one hand, many couples can cope well with the burden of infertility, trying to cope with the treatment, to understand its limitations and to look for alternatives, others are strongly affected by the whole process, so that the management of the case itself is impaired. Thus, the identification of these couples will be of help for treatment.

In general, studies that examine the cognitive response related to infertility tend to perform more evaluations of the general type (Stanton et al., 1991). The psychological reactions to infertility are studied by means of standardized instruments that measure personality, anxiety, depression, and marital life (Newton et al., 1990). However, the use of instruments not specifically designed for infertility involves some drawbacks, even though they offer the advantage of previously determined reliability and validity. One of these drawbacks is the fact that these instruments have been developed for psychiatric use. Thus, the evaluations based on the use of these instruments may contain observations that are interpreted in a different manner when applied to infertile populations.

Based on these considerations, the Centre for Human Reproduction of Fundação Maternidade Sinhá Junqueira, Ribeirão Preto, Brazil, decided to use a psychological evaluation test (PET) consisting of a self-applied questionnaire whose elaboration was based on the repeated complaints of each individual separately, and of couples as a whole. Experience with clinical application of the PET has been reported in previously published papers (Franco Jr et al., 2002a,b).

In terms of structure and operationality, the PET resembles Glover’s adjustment scale (Glover et al., 1999), which is also used to analyse adjustment with a simple and easily applicable questionnaire, which is intended to record individual responses to the stress of infertility. Bernstein’s infertility questionnaire (Bernstein et al., 1985) and Newton’s fertility problem inventory (Newton et al., 1999), although they are also directed at infertility and provide an excellent analysis, have a complex structure, with evaluation divided into sectors. If, on the one hand, they provide a safe measurement of the stress related to infertility, and specific information about different domains of the psychology of the patient, on the other, they involve greater difficulties of application and understanding.

The objective of the present study was to further confirm the ability of the PET to rapidly identify for the clinical staff those patients more affected by infertility who might require psychological support, using an established psychological evaluation method for comparison.

**Materials and methods**

A total of 200 women with an indication for assisted reproduction [specifically IVF/intracytoplasmic sperm injection (ICSI)] attended the Centre for Human Reproduction Fundação Maternidade Sinhá Junqueira were included in the present programme of psychological evaluation by means of the PET. No other type of patient selection was performed. The techniques used for the execution of IVF/ICSI have been described previously (Franco Jr et al., 1995; Mauri et al., 1999).

General patient information concerning age and history of infertility was also recorded.

After the development of the qualitative pilot interviews with the patients (only women) at several stages of investigation, 15 questions were chosen for inclusion in the questionnaire (Table 1). These 15 questions were elaborated on the basis of doubts and/or reports by the assisted reproduction unit patients/couples themselves which were repeatedly presented to the clinical staff. Issues related to infertility identified by other researchers were also considered. The questions were chosen in such a way as to cover variations in cognitive, emotional and behavioural responses in order to provide an indication of the extent to which the women were affected by the impact of infertility (Franco Jr et al., 2002b).

The questionnaires were routinely distributed to the patients, who responded on the occasion of the first control ultrasound after the beginning of ovarian stimulation and before the application of IVF/ICSI. The responses were scored as follows according to intensity (four levels): 1, never or rarely; 2, sometimes; 3, often; 4, always. The sum of the responses was used to determine the final score, which could vary from a minimum of 15 points to a maximum of 60 points. A high PET score was interpreted as representing poor adjustment, with a PET score of ≥30 points being defined as abnormal (cut-off point: median ± 25%). The median was considered, instead of the mean, after the analysis of the curve produced by the scores, which did not follow a normal distribution. To consider the cut-off in median ± 25% meant to consider from 25 to 30% of the studied population as ‘psycho-emotional maladjustment’. The same cut-off was used in a previous paper (Franco Jr et al., 2002b).

The questionnaire was analysed for reliability based on the use of the Cronbach’s alpha coefficient (one name for Cronbach alpha reliability estimate), which measures the extent of each
item on the scale in relation to the other items of the same scale. Cronbach’s alpha is a statistical method that measures the reliability of the tests, observations, experiments or measurements by estimating the extent to which these parameters provide the same results on repeated trials. Cronbach’s alpha can be applied when test items are scored dichotomously, but alpha has the advantage of being applicable when items are weighted (i.e. as in an item scored 1 point, 2 points, 3 points, 4 points). Cronbach’s alpha is a value between 0 and 1 (values near 0 indicate low reliability, values near 1 indicate high reliability).

Simultaneously, the patients were submitted to a psychoanalytical examination by a psychologist who was unaware of the PET results. The clinical method employed, described by Freud, is called preliminary assay, and basically consists of ‘preliminary interviews’ (Freud, 1996; Quinet, 1998). When any decision or proposition of psychoanalytical treatment is to be made, a certain period of time is necessary for observation, to enable better judgement of the case, and to improve the adequacy of treatment. Freud (1996) argued that such preliminary time had the advantage of facilitating diagnosis, and defined it as ‘test treatment’.

Regarding these interviews, Freud stated: “No other type of preliminary test is available, except this procedure, the most extensive debates and questions, in ordinary appointments, do not provide a substitute. This preliminary experiment, however, is itself the beginning of a psychoanalysis and must be according to its rules. Perhaps a distinction can be made, in which the patient is allowed to speak most of time and does not explain more than what is absolutely necessary to enable him/her to continue with what he/she is saying.

There are also diagnostic reasons to start treatment with an experimental period of this type, lasting for 1 or 2 weeks. Often, when a neurosis with hysterical or obsessive symptoms is perceived, which is not excessively conspicuous and has not existed for long, this is exactly the type of case which would be considered appropriate for treatment, considering the possibility that it could be a preliminary stage of what is known as precocious dementia (schizophrenia in Bleuler’s terminology; paraphrenia as I prefer to call it) and that, sooner or later, will present in an advanced stage. I do not agree that it is possible to make the distinction so easily. I am aware there are psychiatrists who hesitate, with less frequency, in their differential diagnosis, but I am convinced that, with the same frequency, they make mistakes. To make a mistake, moreover, is much more serious for the psychoanalyst than for the clinical psychiatrist, as he is called, because the latter is not trying to do something useful, whatever the type of case. He simply takes the risk of making a theoretical mistake and his diagnosis has no more than academic interest. Regarding the psychoanalyst, however, if the outcome of the case is non-favourable, he commits a practical error; he was responsible for unnecessary expenditure and for bringing his method of treatment into disrepute. He cannot follow his promise of cure if the patient is suffering not from hysteria or obsessive neurosis, but from paraphrenia, and therefore the psychoanalyst has particularly strong reasons to avoid making mistakes in diagnosis. In an experimental treatment of some weeks, he will frequently observe suspect signs which will lead him not to proceed further. Regrettably, I cannot assert that such a trial of treatment always prepares us to reach a correct decision; it is simply a wise precaution.”

Today, these interviews have become a consolidated practice in psychoanalysis, representing an initial period for evaluation and potential diagnosis before treatment is started. In practice, the patient is observed over a period of four interviews in a type of study that, although different, maintains the same structure and rules as those of psychoanalysis.

### Table 1. PET for the evaluation of the female population. Distribution of the responses to each item for the 200 patients studied.

<table>
<thead>
<tr>
<th>Questions</th>
<th>1&lt;sup&gt;a&lt;/sup&gt;</th>
<th>2&lt;sup&gt;a&lt;/sup&gt;</th>
<th>3&lt;sup&gt;a&lt;/sup&gt;</th>
<th>4&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Are you irritated by the fact of not having children?</td>
<td>69</td>
<td>77</td>
<td>35</td>
<td>19</td>
</tr>
<tr>
<td>2 Relatives and friends usually ask about the fact that we don’t have children and I don’t feel comfortable in this situation</td>
<td>47</td>
<td>75</td>
<td>43</td>
<td>35</td>
</tr>
<tr>
<td>3 I am upset when I am invited to a children’s birthday party</td>
<td>146</td>
<td>38</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>4 I am annoyed when a friend or relative becomes pregnant</td>
<td>104</td>
<td>54</td>
<td>19</td>
<td>23</td>
</tr>
<tr>
<td>5 Are you depressed each time you menstruate?</td>
<td>50</td>
<td>77</td>
<td>36</td>
<td>37</td>
</tr>
<tr>
<td>6 Is your sexual relationship being impaired by the fact that you have not become pregnant up to now?</td>
<td>144</td>
<td>43</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>7 Is your professional activity being impaired due to the lack of children?</td>
<td>164</td>
<td>21</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>8 Do you feel inferior to other women due to the fact of not having children?</td>
<td>98</td>
<td>63</td>
<td>22</td>
<td>17</td>
</tr>
<tr>
<td>9 Are you a person who is always suspicious or afraid of treatments?</td>
<td>117</td>
<td>67</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>10 Do you think you might go crazy if you don’t have children?</td>
<td>164</td>
<td>26</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>11 Do you feel tachycardia, shortness of breath, pressure in the chest, tremors, and hand sweating when thinking about the fact of not having children?</td>
<td>139</td>
<td>44</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>12 Do you feel a sensation of emptiness due to the fact of not having children?</td>
<td>43</td>
<td>86</td>
<td>37</td>
<td>34</td>
</tr>
<tr>
<td>13 Is your daily relationship with your husband impaired by the fact of not having children?</td>
<td>138</td>
<td>44</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>14 Does the difficulty in having children make you want not to leave home as you used to do and to think that it is better to be isolated from others?</td>
<td>163</td>
<td>22</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>15 Do you think about your difficulty in having children during daily life?</td>
<td>31</td>
<td>82</td>
<td>40</td>
<td>47</td>
</tr>
</tbody>
</table>

<sup>a</sup>Key: (1) never or rarely; (2) sometimes; (3) often; (4) always.
The functions of preliminary interviews are of three types, with a distribution that is logical rather than chronological. First, the function ‘of the symptom’, when the symptom is questioned by the psychologist, who will try to find out to what the symptom corresponds. Second, a diagnostic function is used to guide the analysis, permitting a differential and structural diagnosis. Finally, a transference function is necessary for the actions of analysis to be applied. From the viewpoint of the psychologist, the preliminary interviews may be divided into two stages: a time to understand and a time to conclude. It is at this time of conclusion that the psychoanalytical act occurs to transform treatment from an essay to analysis proper.

When the psychologist, after the four interviews, concluded that the patient required continued psychological follow-up, the evaluation was considered to be abnormal. If, on the contrary, no additional procedure was indicated, the evaluation was considered to be normal.

Epidemiological tests were used for statistical analysis, with the determination of efficiency (the extent to which a test produces a result in accordance with the gold standard), sensitivity (the proportion of those with the disease or condition, as measured by the gold standard, who are positive by the test being studied, synonym: positive-in-disease), positive predictive value (the proportion of people with a positive test who actually have the condition or disease as measured by the gold standard, synonym: post-test probability after a positive test), specificity (the proportion of those without the disease or condition, as measured by the gold standard, who are negative by the test being studied, synonym: negative-in-health), negative predictive value (the proportion of individuals with a negative test do not have the condition or disease as measured by the gold standard, synonym: post-test probability after a negative test), likelihood ratio for a positive test (a ratio of the probability of a positive test if the disease is present to the probability of a positive test if the disease is absent), and likelihood ratio for a negative test (a ratio of the probability of a negative test if the disease is present to the probability of a negative test if the disease is absent). Data were analysed statistically using the InStat 3.0 software for MacIntosh (GraphPad Software, San Diego, CA, USA).

**Results**

All the 200 women included in the study responded to the PET and completed the four psychological interviews. The analysis of the general data is presented in Table 2.

The internal consistency of all the scale items was tested using the Chronbach alpha coefficient. The general coefficient of the test was 0.88, showing that the items were highly correlated, since the exclusion of any question did not lead to a drastic change in the alpha value. A level >0.75 is usually considered acceptable.

The scores ranged from a minimum of 15 to a maximum of 57 points. Table 1 shows the distribution of the responses (1–4) for each question formulated.

The PET scores were compared with the results of psychological examination. In the PET result, 66 of the 200 patients (33%) scored a total of 30 points or more and 134 (67%) scored less than 30 points. In the preliminary interviews, 105 (52.5%) patients showed an abnormal evaluation and 95 (47.5%) a normal evaluation. The correlation between PET and psychological evaluation is demonstrated in Table 3.

Statistical analysis of the data showed 82% efficiency, 62% sensitivity, 98% positive predictive value, 98% specificity, 70% negative predictive value, likelihood ratio for a positive test 62, and likelihood ratio for a negative test 0.38.

**Discussion**

The PET is a diagnostic test and should be analysed as such. Although diagnostic tests are traditionally viewed as a means of reducing the diagnostic doubts, they can be used only if the clinicians understand how these tests reduce the uncertainties about the condition of the patient and how the tests describe and quantify them.

The fundamental principle of tests in general is based on the belief that individuals ‘with a disease’ are different from ‘healthy’ individuals and that the test will be able to distinguish between these two groups. In an ideal situation, a diagnostic test has the following characteristics: (i) all individuals ‘without a disease’ under study show a uniform value for the test, (ii) all individuals ‘with the disease’ under study show a uniform value for the test but different from the previous one, and (iii) all the results of the tests will coincide with the results for the group with the disease or for the group without the disease. If this situation were real, then a perfect

<table>
<thead>
<tr>
<th>Table 2. General data and aetiology for the study population.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age (years) ± SD (range)</td>
</tr>
<tr>
<td>Primary infertility (%)</td>
</tr>
<tr>
<td>Secondary infertility (%)</td>
</tr>
<tr>
<td>Male factor (n)</td>
</tr>
<tr>
<td>Unexplained infertility (n)</td>
</tr>
<tr>
<td>Peritoneal tube (n)</td>
</tr>
<tr>
<td>Endometriosis (n)</td>
</tr>
<tr>
<td>Male + tubal (n)</td>
</tr>
<tr>
<td>Male + endometriosis (n)</td>
</tr>
<tr>
<td>Endometriosis + tubal (n)</td>
</tr>
</tbody>
</table>

Values in parentheses are percentages unless otherwise stated.
aNumber of patients.

<table>
<thead>
<tr>
<th>Table 3. Comparison of the psychological evaluation test (PET) scores and results at preliminary interviews.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PET score</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>≥30</td>
</tr>
<tr>
<td>≤30</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>
test could distinguish between health and disease and the task of the clinician would simply be to prescribe it. However, what commonly occurs is the presence of variations in each of the three basic factors.

Since the PET is not a standardized test and since some authors have highlighted limitations of psychological evaluations of the questionnaire type (Fassino et al., 2002), it was decided to study the ability of the test to select patients who require psychological support by comparing it to a classical type of psychological evaluation (preliminary interviews). Other tests focusing on infertility questions (Bernstein’s infertility questionnaire, Newton’s fertility problem inventory, Glover’s adjustment scale), or even tests of more general observation could be used for comparison. However, since the use of questionnaires per se may induce methodological bias (reliability of the response), it was preferable to use for comparison a more general evaluation of the patients, of possible application but close to normal psychological follow-up. In studies in which a test specifically elaborated for the evaluation of the effects/result of infertility was used (Bernstein et al. 1985; Glover et al., 1999; Newton et al., 1999), when the test was compared with other evaluations, general existing tests were used rather than an individual examination by a psychologist. The results obtained showed that the PET score had a high capacity to distinguish between affected and healthy patients (82% efficiency of the test), with a highly precise selection of those patients most affected by emotional changes (99% specificity, 98% positive predictive value). In addition, patients with an abnormal PET result showed a 62-fold higher probability of requiring psychological support (likelihood ratio for a positive test 62).

On the basis of its form of execution (self-filled questionnaire), the PET was indeed an instrument for the rapid identification of patients who might benefit from more specific and direct psychological support.

One point that might be criticised is the number of patients who presented emotional changes on the basis of psychological examination but were not detected by the PET (62% sensitivity, negative predictive value and likelihood ratio for a negative test 0.38). However, this point deserves some consideration.

First, the PET was not developed as the only instrument for psychological evaluation of a woman. It is simply a test that has its place and utility within a context, as also recommended by Bernstein (1985). Women are followed up in a continuous manner and those that are missed due to any methodological bias of the PET or who will present emotional changes with time and with the development of a history of infertility could be later referred for psychological care.

Second, more thought should be addressed to the objectives of a test. This is not a ‘high morbidity–lethality disease’ which requires a test capable of identifying every individual with the possibility of being involved, even when this is not confirmed later (high sensitivity). On the other hand, in psychology success largely depends on the awareness of the individual, that he or she needs support. Several studies have shown that, despite the unanimous opinion about the need for and the potential beneficial effect of psychological care, few patients use this service when it is available (Laffont and Edelmann, 1994; Boivin et al., 1999). Thus, it may be worse to refer individuals for psychological care when there is no need for it (false-positive results, very low in the PET) than to miss some individuals (false-negative results), who might be identified later.

A third point is the validity of identifying affected patients when the prevalence of the alteration is more than 50% of the population involved. However, it is difficult to define levels of such expression as the consequence of an anguishing condition experienced by those who are unable to conceive a child, with many manifestations remaining at the subclinical level. The basic usefulness of the PET is not the identification of these individuals, but rather the identification of those most severely involved, i.e. those who require a stronger intervention for the relief of emotional stress.

In conclusion, the PET proved to be a useful clinical instrument, being of help in the selection of patients examined for the possible presence of psychological need. As a clinical mechanism, the test can facilitate the beginning of a discussion on adjustment to infertility problems with the couples.

References


Domar AD, Broome A, Zuttermeister PC et al. 1992 The prevalence and predictability of depression in infertile women. *Fertility and...
Sterility 58, 1158–1163.
King RB 2003 Subfecundity and anxiety in a nationally representative sample. Social Science and Medicine 56, 739–751.

Received 4 August 2003; refereed 29 August 2003; accepted 30 September 2003.