PROGNOSTIC VALUE OF THE POSTCOITAL TEST FOR SPONTANEOUS PREGNANCY

Esther Leushuis
Jan Willem van der Steeg
Pieterneel Steures
Carolien Koks
Jur Oosterhuis
Petra Bourdrez
Patrick M.M. Bossuyt
Fulco van der Veen
Ben Willem J. Mol
Peter G.A. Hompes for CECERM

FERTILITY AND STERILITY
2011 MAY;95(6):2050-6
ABSTRACT

OBJECTIVE
To evaluate the capacity of the postcoital test (PCT) to predict spontaneous pregnancy in a large cohort study of subfertile couples.

DESIGN
Prospective study.

SETTING
Department of reproductive medicine of 38 hospitals in The Netherlands.

PATIENTS
Between January 2002 and February 2004, we prospectively included consecutive subfertile couples who had not been evaluated previously for subfertility.

INTERVENTIONS
We estimated the contribution of the PCT result to the existing prediction model for spontaneous pregnancy by calculating the adjusted hazard ratio (HR) of an abnormal PCT result. We constructed a second prediction model (PCT model) based on the reference model including the PCT.

MAIN OUTCOME MEASURE(S)
Primary endpoint in this study was ongoing pregnancy. We evaluated the performance of the PCT model in comparison with the reference model by calculating goodness of fit, discrimination, calibration, and the “net reclassification improvement”.

RESULTS
We included 3,021 couples of whom 537 (18%) had a spontaneous pregnancy and 55 (1.8%) a nonsuccessful pregnancy; 1,316 (44%) started treatment within 12 months, 824 (27%) neither started treatment nor became pregnant and 289 (10%) became lost to follow-up within 12 months. The adjusted HR for an abnormal PCT was 0.76 (95% confidence interval [CI]: 0.62 to 0.94). The adjusted HR for an abnormal PCT was 0.63 (95% CI: 0.47 to 0.84) in case of no spermatozoa, 0.81 (95% CI: 0.57 to 1.2) in case of nonmotile spermatozoa and 1.2 (95% CI: 0.8 to 1.8) in case of motile, nonprogressive spermatozoa. Adding PCT to the reference model did not significantly improve goodness of fit. Discrimination was equally poor for the PCT model and the reference model. The calibration plots of both models showed comparably good calibration. The net reclassification improvement of the predictions of the PCT model compared with the reference model was −1.1%.

CONCLUSION
This study demonstrated that the postcoital test has prognostic value but does not add substantially to a prognostic model for spontaneous pregnancy.
INTRODUCTION

Since the original description of the postcoital test (PCT) in 1866 by J. Marion Sims \(^1\) and the reintroduction by M. Huhner in 1913 \(^2\) its use has become widespread in the evaluation of subfertile couples. Many authorities considered the PCT as the cornerstone of the fertility evaluation. \(^3\) Although the PCT is used commonly, the use of this test in the basic fertility workup has been subject to debate over the last 10 years due to conflicting data.\(^4\)-\(^6\)

On the one hand, the PCT enables identification of couples with a cervical factor and avoids misclassifying these couples as having unexplained infertility, which has important implications for treatment. An abnormal PCT result decreases the probability of treatment independent pregnancy twofold to threefold.\(^5\)-\(^9\) On the other hand, routine use of the PCT would only lead to more interventions without an increase in pregnancy rates.\(^10\)

Recently, we validated a prediction model for spontaneous pregnancy in a large cohort of subfertile couples.\(^11\) In this validation, the prediction model including female age, duration of subfertility, primary subfertility, sperm motility and referral status had good predictive performance, but adding PCT results to this model did not improve its performance.

Thus, at present, data on the prognostic value of the PCT for spontaneous pregnancy in subfertile couples are inconclusive. The aim of this study was to evaluate the prognostic value of the PCT in a large prospective multicenter cohort of subfertile couples in relation to an existing, well validated prediction model for spontaneous pregnancy.

MATERIALS AND METHODS

Validation cohort

The study was designed as a prospective cohort study performed in 38 hospitals in the Netherlands. The Institutional Review Board of the Academic Medical Centre approved the study, and local approval was obtained from the Board of Directors from each participating hospital. Between January 2002 and February 2004, we included consecutive subfertile couples who had not been evaluated previously for subfertility.\(^11\) In couples who had an unfulfilled wish for a child and had 1 year with regular unprotected intercourse and in whom the woman had a regular cycle, a basic fertility workup was performed, consisting of a fertility history, semen analysis, PCT, assessment of ovulation, and assessment of the Fallopian tubes according to the guidelines of the Dutch Society of Obstetrics and Gynaecology (2004).

The performance of the PCT was standardized as described in detail in the study protocol for all participating centers. At least one PCT was performed during the infertility workup.\(^5\)-\(^9\),\(^11\) The PCT could be planned on the basis of the basal body temperature and cycle length or in repetitive ultrasound findings. The PCT was judged to be normal if at least one progressively motile spermatozoon was seen in one of five high-power fields at \(\times400\) magnification. The PCT was judged abnormal if motile, nonprogressive spermatozoa or nonmotile or no spermatozoa were seen in one of the five high-power fields at \(\times400\) magnification.
Follow-up
The follow-up of the couples is described in full detail in the publication of van der Steeg et al. Primary endpoint was spontaneous conception resulting in an ongoing pregnancy. Spontaneous ongoing pregnancy was defined as the presence of fetal cardiac activity at transvaginal sonography at a gestational age of at least 12 weeks, resulting from a treatment-independent conception. Time to pregnancy was considered censored at the moment treatment had been started or at the last date of contact during follow-up, when the couple had no ongoing pregnancy. For all couples lost to follow-up, the general practitioner was sent a questionnaire on the most recent fertility status of the couple.

Data analysis
Missing data for the predictive variables were imputed (‘filled in’), because deleting them would lead to a loss of statistical power and potentially biased results in multivariable analysis. The technique of imputation and details of this specific dataset are described in a previous publication. The Hunault model for predicting spontaneous pregnancy was considered the reference model. This model includes five prognostic variables: female age, duration of subfertility, female subfertility being primary or secondary, percentage motile spermatozoa of the first semen analysis, and referral status. Referral by a gynaecologist was established as a confounding factor in the original publication. We then evaluated the additional value of the PCT by building a new multivariable Cox model: the PCT model. This PCT model includes the variables from the reference model and, additionally, the PCT outcome. We did not recalculate the weights for the variables in the reference model, as this would lead to a loss in power and an increased risk of capitalization on chance: an artificial increase in predictive power which is due only to chance fluctuations in the variables. We therefore used the linear combination of all variables in the reference model. In other words, for each couple we calculated the sum of the variables in the reference model, each variable weighted by the respective coefficient as estimated by Hunault et al.. We then estimated two coefficients for our PCT model using multivariable Cox regression analysis: one coefficient for the sum of the weighted variables of the reference model, and a second coefficient for the PCT outcome. We transformed the coefficients into hazard ratios (HR) to facilitate interpretation. We performed a separate analysis to assess the prognostic value of each separate abnormal PCT category relative to a normal PCT: motile, nonprogressive spermatozoa and nonmotile or no spermatozoa.

Evaluation of performance
We compared the goodness of fit of the PCT model to our data to the goodness of fit of the reference model using the generalized likelihood ratio test. This test evaluates whether adding the PCT to the reference model leads to a significant improvement in goodness of fit. We then used three concepts to compare the performance of the PCT model with that of the reference model. First, we evaluated the discriminative power of the models. Discrimination refers to the ability to distinguish couples who will conceive from those who will not. Second, we evaluated the calibration of the
models. Calibration is the level of correspondence between the calculated pregnancy probabilities and the observed proportion of pregnancies. For a comprehensive explanation of discrimination and calibration and the interpretation of performance we reference our recent publications.\textsuperscript{11,15} Third, we determined the degree of reclassification between the reference model and the PCT model by calculating the net reclassification improvement (NRI).\textsuperscript{16,17} This method is based on the difference between two models in the individual calculated probabilities of a pregnancy and the occurrence of a pregnancy in couples with probabilities that are discordant between the two models. Calculating the net reclassification improvement requires a priori meaningful risk categories. We used 0\% to 30\%, 30\% to 40\%, and >40\% for the probability of spontaneous pregnancy.\textsuperscript{11} In net reclassification improvement, only those changes in estimated probabilities that imply a change from one category to another (reclassification) are considered. The reclassification of couples who develop and who do not develop a pregnancy should be considered separately. Any “upward” movement in categories for couples who achieved a pregnancy implies improved classification, and any “downward movement” indicates worse reclassification. The interpretation was opposite for couples who did not achieve a pregnancy. The net improvement in reclassification was quantified as a sum of differences in proportions of couples moving up minus the proportion moving down for couples who develop events, and the proportion of couples moving down minus the proportion moving up for couples who do not develop events.

RESULTS

We completed the fertility workup of 7,860 couples. Of these, 948 had a severe male factor, 311 had two-sided tubal pathology and 1,311 had one-sided tubal pathology, and 2,642 had other reasons for exclusion (i.e. prior fertility treatment), leaving 3,021 couples for inclusion. The baseline characteristics of these couples are presented in Table 1. We were able to complete follow-up for 2,741 couples (90\%). Of the 3,021 couples, 537 (18\%) had a spontaneous ongoing pregnancy within 1 year, including 10 multiple pregnancies (0.3\%). Within 12 months 1,316 (44\%) of all couples had started treatment, whereas 824 (27\%) neither had started treatment yet nor had become pregnant within the follow up period. The follow-up status of all patients at 12 months is shown in Figure 1.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean or Median</th>
<th>5th-95th percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female age (y)</td>
<td>32.5</td>
<td>25 – 39</td>
</tr>
<tr>
<td>Male age (y)</td>
<td>34.8</td>
<td>27 – 44</td>
</tr>
<tr>
<td>Duration of subfertility (y) (median)</td>
<td>1.7</td>
<td>1.0 – 3.9</td>
</tr>
<tr>
<td>Subfertility, primary (n)</td>
<td>2,013</td>
<td>67%</td>
</tr>
<tr>
<td>Sperm motility (grade A WHO %)</td>
<td>42.4</td>
<td>10 – 75</td>
</tr>
<tr>
<td>PCT\textsuperscript{a}, normal (n)</td>
<td>2,050</td>
<td>68%</td>
</tr>
</tbody>
</table>

\textit{WHO = World Health Organization. a Postcoital test.}
The mean probability of a spontaneous pregnancy as calculated with the reference Hunault model without PCT was 0.32 (5th and 95th percentiles: 0.16 and 0.52). For the PCT model, the mean calculated probability was 0.29 (5th and 95th percentiles: 0.16 and 0.45).

The adjusted HR of an abnormal PCT result in the PCT model was 0.76 (95% confidence interval [CI]: 0.62 to 0.94). The adjusted HR of an abnormal PCT was 0.63 (95% CI: 0.47 to 0.84) in case of no spermatozoa, 0.81 (95% CI: 0.57 to 1.2) in case of nonmotile spermatozoa and 1.2 (95% CI: 0.8 to 1.8) in case of motile, nonprogressive spermatozoa, each relative to a normal PCT.

**Evaluation of performance**

Goodness of fit was not improved significantly when the PCT model was compared with the reference model (difference in -2 log likelihood: 7, P>.05 [degrees of freedom 3]). Discrimination was comparable for the PCT model (area under the curve: 0.64; 95% CI: 0.61 to 0.66) and the reference model (area under the curve: 0.63; 95% CI: 0.60 to 0.65), a nonsignificant difference. The calibration plots of both models showed good calibration (Fig. 2). The calibration of the PCT model was slightly better than the
calibration of the reference model.

The addition of the PCT to the reference model reduced the correct classification of couples who achieved a pregnancy with 19.0%, but improved the correct classification of the couples who did not achieve a pregnancy with 17.9%. The “net reclassification improvement” of the predictions of the PCT model compared with the reference model was –1.1%. The PCT model showed a minor net reduction in the correct classification of couples who did and did not achieve a pregnancy compared to the reference model. The details of the classification per model per couple are shown in table in Table 2.

DISCUSSION

This study showed that subfertile couples with an abnormal PCT result have a lower probability of a spontaneous pregnancy, especially if no sperm are found in the cervical mucus. Despite the prognostic information of the PCT, addition of the PCT outcome as a prognostic variable to a validated prediction model for spontaneous pregnancy did not improve the fit of this model significantly and even led to a nonsignificant reduction in the correct classification of couples who did and did not achieve a pregnancy.

The prognostic value of the PCT for spontaneous pregnancy has so far been assessed in multivariable models and in one randomised controlled trial.5;7;9;11;14;18;19 None of these studies was designed specifically to assess the prognostic value of the PCT to the

Table 2. Net reclassification improvement of the PCT model compared with the reference model.

<table>
<thead>
<tr>
<th>PCT model, frequency (row percent)</th>
<th>&lt; 30%, n (%)</th>
<th>30% to 40%</th>
<th>&gt; 40%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference model</td>
<td>No.</td>
<td>%</td>
<td>No.</td>
</tr>
<tr>
<td>Couples who achieved a pregnancy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;30%</td>
<td>179</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>30% to 40%</td>
<td>41b</td>
<td>16</td>
<td>149</td>
</tr>
<tr>
<td>&gt;40%</td>
<td>0b</td>
<td>0</td>
<td>61b</td>
</tr>
<tr>
<td>Total</td>
<td>220</td>
<td>40</td>
<td>210</td>
</tr>
<tr>
<td>Couples who did not achieve a pregnancy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;30%</td>
<td>1,113</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>30% to 40%</td>
<td>228b</td>
<td>24</td>
<td>492</td>
</tr>
<tr>
<td>&gt;40%</td>
<td>0b</td>
<td>0</td>
<td>169</td>
</tr>
<tr>
<td>Total</td>
<td>1,341</td>
<td>58</td>
<td>661</td>
</tr>
</tbody>
</table>

Note: The net reclassification for couples who achieved a pregnancy is a deterioration of 19% and for the couples who did not achieve a pregnancy is an improvement of 17.9%. This results in an overall net reclassification improvement of the PCT model of -1.1%.

a Deterioration of classification for the PCT model.
b Improvement of classification for the PCT model.
already established prognostic factors for spontaneous pregnancy. The first study of 726 subfertile couples presented an adjusted HR for an abnormal PCT of 0.26 (95% CI 0.17 to 0.40) for the prediction of live birth, but two out of four variables selected for the final model—tubal defects and anovulation—should have been excluded because of the manifestly reduced chances of spontaneous conception in couples with these factors. Moreover, the HR of the PCT in this study could have been overestimated because of insufficient adjustment for the established prognostic variables.

A second study, in 996 subfertile couples, reported that a progressive PCT was the most important prognostic variable of all, with a HR of 0.23 (95% CI: 0.14 to 0.4) 5, but the model derived from this study demonstrated poor calibration at external validation. The model underestimated the pregnancy rates in couples with a poor prognosis (0% to 20%) and overestimated them in couples with a good prognosis (≥40%). A third study was a randomized clinical trial. The authors of this trial concluded that the routine use of the PCT leads to more interventions without an associated increase in pregnancy rates. This study was criticized because a substantial number of participating women were anovulatory and because no specific recommendations were made on how to manage a couple with an abnormal PCT. A fourth study in 207 couples evaluated the prognostic value of the PCT in a multivariable model, including the variables duration of subfertility, age, PCT result, and motile sperm concentration. Surprisingly, this study showed no prognostic value for woman’s age or motile sperm concentration. The added prognostic value of the PCT, relative to the established factors for spontaneous pregnancy could not be assessed, possibly because of the size of the study population but primarily because of a lack of complete data to enable assessment of all prognostic factors and model performance. The last study to evaluate the PCT outcome together with the established factors generated two models for the prediction of spontaneous pregnancy, one with and one without the PCT, showing inferior performance for the model without the PCT. However, the derivation of the model with the PCT was based on a smaller cohort of patients (n = 1,398) than the models without the PCT (n = 2,459), which could have led to a less accurate estimation of regression coefficients. Further, these models were validated in a prospective validation cohort study. The performance of the model with the PCT was inferior to the model without the PCT, but the primary aim of that study was to evaluate the performance of the model without the PCT, and the techniques used in the external validation for the assessment of performance were limited.

One of the strengths of our study is that we were able to assess the specific value of the PCT in addition to the already established factors for the prediction of spontaneous pregnancy in a large prospective multicenter cohort of subfertile couples. We excluded couples with known causes of subfertility that manifestly reduce chances of spontaneous conception. We were able to include a large set of variables and used the firm evidence gained from an established prediction model to build onto. The statistical analysis of this study allowed for the time of spontaneous chance to conception by censoring the outcomes, even if the couples had undergone treatment within 1 year. In an ideal world it would have been better to allow all couples the
opportunity of spontaneous conception over the same time period; nowadays this is difficult to achieve. The outcome of couples that were lost to follow-up can be regarded as being missing completely at random. There is no reason to assume that positive results are systematically more often missing than negative results. The couples that were lost to follow-up were classified as not becoming pregnant.
The PCT allows identification of couples where the spermatozoa do not penetrate the cervical mucus. In a previous randomized clinical trial, we showed that, in couples with an abnormal PCT but a good prognosis, IUI in the unstimulated cycles was superior over no treatment. Similarly, in couples with an abnormal PCT but otherwise also a poor prognosis, IUI without controlled ovarian hyperstimulation (COH) leads to pregnancy rates that are comparable with those obtained by IUI with COH. IUI without COH should therefore be the treatment of first choice in these couples, as it reduces costs and the risk of multiple pregnancies, without compromising pregnancy rates.

Although we found that the PCT does not help in establishing a prognosis for spontaneous conception, it might identify couples who do not conceive as a consequence of sexual dysfunction, who may benefit from sexual therapy or insemination in the natural cycle.

If we abandon the PCT, we should realize that we potentially do not identify some couples with subfertility who actually have sexual dysfunction. The PCT is the only test in the fertility workup that endeavours to assess sexual function in vivo, thus allowing identification of subfertile couples who do not conceive because they fail to have successful intercourse, whereas their biological components necessary for conception—semen, ovulation, and tubal function—are all normal. Studies reporting on the PCT showed that in 7% to 16% no spermatozoa were seen. In these couples no pregnancies were observed. No spermatozoa were seen in 17% of all PCTs in our study, but 49 pregnancies were seen. The PCT or a variant might be part of the fertility workup in a research setting.

In summary, our study demonstrated that the postcoital test has prognostic value but does not significantly add value to a prognostic model for spontaneous pregnancy. This adds to a previous randomized clinical trial that indicated no benefit of the PCT. The test adds no prognostic information to existing models, and the predictive information from the PCT is already incorporated in other variables based on the semen analysis.
REFERENCES

20. Bossuyt PM, Lijmer JG, Mol BW. Randomised comparisons of medical tests: sometimes invalid,


CHAPTER 7
PROGNOSTIC VALUE OF THE POSTCOITAL TEST FOR SPONTANEOUS PREGNANCY