Diagnosing tubal pathology: The individual approach

Broeze, K.A.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 7

Tubal patency tests and predicting fertility outcome. An individual patient data meta-analysis

Kimiko Broeze
Femke van Zanten
Patrick Bossuyt
Brent Opmeer
Nan van Geloven
John Collins
Carolien Koks
Lena Lindborg
Piotr Marianowski
Denise Perquin
Jan Willem van der Steeg
Pieternel Steures
Fulco van der Veen
Ben Willem Mol

Fertility Sterility 2013. Submitted
Chapter 7

Abstract

**Introduction** To provide optimal management for subfertile couples, it is essential to calculate chances of natural conception. The aim of this study was to investigate the added value of tubal patency tests over patient characteristics and to assess the predictive capacity of these tests for natural conception in women with different profiles.

**Material & Methods** Authors of selected articles were approached to share their data. Models were built to assess the predictive value of patient characteristics and the added value of HSG and/or laparoscopy. Fecundity rate ratios (FRRs) were calculated for women with specific profiles.

**Results** We obtained data of 6,326 women. The model based on female age, duration and type of subfertility, BMI, previous Chlamydia infection and a positive CAT had an area under the ROC curve of 0.61 (95% CI 0.60 - 0.63). Adding HSG, laparoscopy and both HSG and laparoscopy, increased the AUCs to 0.63 (95% CI 0.62 - 0.65), 0.62 (95% CI 0.60 - 0.63) and 0.63 (95% CI 0.62 - 0.65) respectively. Higher female age and longer duration of subfertility lead to lower rates of natural conception.

**Conclusion** Natural conception can be best predicted from patient characteristics and HSG results. Laparoscopy had no added value to HSG.
Introduction

Subfertility affects 10% to 15% of couples trying to conceive (Brandes et al., 2010). One of the major causes of female subfertility is tubal pathology, of which the assessment by hysterosalpingography (HSG), hystero-contrast-sonography (HyCoSy), transvaginal hydrolaparoscopy (THL) or diagnostic laparoscopy (DLS) plays an important role in the diagnostic work-up of subfertile couples (ASRM, 2006; NICE, 2004; NVOG, 2004). This work-up is helpful in judging the likelihood of natural conception. A high probability of natural conception justifies an expectant management policy, whereas starting assisted reproductive treatment can be offered if the probability for natural conception is less than 30% (Steures et al., 2006).

Previous studies have shown that tubal pathology detected by HSG or laparoscopy affects fertility outcome. In case of unilateral tubal pathology at HSG or DLS, there was a slight reduction in spontaneous pregnancy chances, with bilateral tubal pathology, spontaneous pregnancy rates were more strongly affected (Nordenskjold and Ahlgren, 1983; Mol et al., 1997; Mol et al., 1999; Mol et al., 1999; Verhoeve et al., 2011).

All previous studies estimated the predictive capacity of HSG and laparoscopy as stand alone modalities, without taking into account patient characteristics like female age, duration of subfertility, previous pelvic inflammatory disease (PID) or previously performed pelvic surgery. Thereby these evaluations ignored possible relations between patient profile and tubal status. It is still unknown to what extent HSG and laparoscopy have added value in the prediction of spontaneous pregnancy on top of the information already known from clinical history. Also, the predictive value of HSG and laparoscopy might vary for patients with different profiles. It is possible that for some patients a work-up with HSG might be sufficient, whereas others could start immediately with laparoscopy.

The aim of this study was to investigate the possible added value of these tests in the prediction of natural conception, relative to patient characteristics obtained from clinical history and physical examination, and to reassess the predictive capacity of tubal patency tests for spontaneous pregnancy chances in women with different prognostic profiles.

Methods

Literature search

A computerized search in Medline and Embase was performed to identify registered papers published between January 1995 and December 2010. Keywords used were ‘female subfertility’ AND ‘hysterosalpingography’ OR ‘hysterosalpingo-contrast-sonography’ OR ‘transvaginal hydrolaparoscopy’ OR ‘diagnostic laparoscopy’ AND ‘pregnancy’. Studies were considered eligible if the included women had been referred for tubal patency testing by HSG or laparoscopy and if pregnancy outcome was reported. Besides HSG and laparoscopy
we also included studies of hysterosalpingo-contrast-sonography (HyCoSy) or transvaginal hydrolaparoscopy (THL), as the findings from these tests correspond to a large extent with those from HSG (Ekerhovd et al., 2004) and laparoscopy (Fujiwara et al., 2003), respectively. Studies were selected in a two-stage process. First, two reviewers scrutinized titles and abstracts from the electronic searches independently (KB and FvZ) and full manuscripts of all citations that were likely to meet the predefined selection criteria were obtained. Final inclusion or exclusion decisions were then made on examination of the full manuscripts. Differences of opinion were resolved by consensus after consultation with a third researcher. References of obtained articles were scanned to identify other potentially eligible articles. Authors of eligible articles were asked whether they were aware of any additional eligible studies. This way, data from studies missed by our search criteria, or from studies that have not been published at all, were also eligible for inclusion. The search was not restricted by language.

Data acquisition

For each eligible article, we obtained contact information on the first, second or last author. In case contact information on the first author was not available or if the first author did not respond, we contacted the second or last author. We approached these authors by mail and invited them to share their data in this collaborative project. Authors willing to participate were asked to send their original dataset. We requested the complete database in the original format, as to minimise authors’ efforts in selecting the appropriate variables or in converting data to a specific format. If variables and categories were not adequately labelled within the dataset, a clear legend in English was asked for.

Datasets should at least include the following variables: anonymous patient identifier, data on patient characteristics such as female age or duration of subfertility, results of HSG, HyCoSy, THL or diagnostic laparoscopy (tubal pathology absent or present), spontaneous pregnancy or live birth and time to pregnancy. Pregnancy was defined as an ongoing pregnancy at a gestational age of 12 weeks, confirmed by ultrasonography.

The present analysis was limited to couples with a regular ovulatory cycle, defined as a cycle length between 23 and 35 days with a within cycle variation of less than eight days. Duration of subfertility was defined as the time between child wish and performance of the tubal patency test of interest, or, in the studies that included two tests, as the time between child wish and performance of the first performed test. Type of subfertility was defined as primary subfertility when a couple had no previous pregnancies or secondary infertility if a couple had a previous pregnancy. Previous pelvic inflammatory disease (PID) was defined as a clinical episode of abdominal pain and fever for which antibiotics were given and no other focus was found. Previous Chlamydia infection was defined by a positive PCR for Chlamydia without clinical signs of infection, or by self-reporting of the patient. The following Chlamydia antibody tests were used: ELISA (cut-off 1.1), IF (cut-off 1:32) and MIF
(cut-off 1:32). Time to pregnancy was defined as the time between the tubal patency test of interest and the occurrence of a spontaneous ongoing pregnancy. Tubal pathology included unilateral and bilateral tubal pathology. At HSG or HyCoSy, tubal pathology was defined as impaired or absent flow of contrast medium in the fallopian tubes, either proximally or distally. At TLH or laparoscopy, tubal pathology was defined as the presence of occlusion of the fallopian tubes, with or without hydrosalpinges or peritubal adhesions. Tubal occlusion was considered to be present if filling or spillage of dye was not observed. Unilateral tubal pathology was defined as the presence of occlusion of the fallopian tubes, with or without hydrosalpinges or peritubal adhesions, in one of the tubes. Bilateral tubal pathology was present when such abnormalities were seen in both tubes. If needed, ethical approval had been acquired by the original authors.

**Quality assessment**

The quality of every included study was scored with a modified QUADAS tool by two researchers independently (KB and FvZ). The QUADAS checklist has been developed for use in systematic reviews of diagnostic accuracy studies. We modified this checklist to fit the quality assessment of prognostic studies (Whiting et al., 2003). Additional items were created for the description of selection criteria, execution of index tests (HSG and laparoscopy) and the description of follow-up (duration of follow-up, as well as loss to follow-up).

We used RevMan 5 software (Cochrane Collaboration) to summarize the quality indicators of the included studies. Completeness of the datasets was described, based on the availability of data on patient identifiers, diagnostic test results and pregnancy rates. We compared the acquired data and the published results for consistency. We also checked the included studies for their study characteristics, study design, inclusion criteria and diagnostic strategy. Participating authors were contacted to confirm missing data or to discuss major discordant results between acquired data and reported data.

**Statistical analyses**

We merged all acquired original data into a summary database if variables were compatible. Incompatible data were recoded and also added to the summary database. Time to pregnancy was calculated as the time between test performance and the occurrence of pregnancy or considered censored at the start of assisted reproductive treatment (OI, IUI, IVF or ICSI) or at the end of follow-up if no pregnancy occurred. Pregnancies after ovulation induction (OI), intra-uterine insertion (IUI), in vitro fertilisation (IVF or ICSI) or tubal surgery were excluded.

For all analyses, results of HSG and HyCoSy were analysed as one variable, as well as results of THL and laparoscopy, since data on HyCoSy and THL were scarce and these tests correspond to a large extent with HSG (Ekerhovd et al., 2004) and laparoscopy (Fujiwara et al., 2003), respectively.
First, we performed a missing value analysis and used multiple imputation to allow inclusion of observations with missing patient characteristics or missing results of HSG and laparoscopies. We conducted five imputation rounds, in which all available variables, including the diagnostic test results, ongoing spontaneous pregnancy and the individual study identifier were used as predictors to optimise the imputation (Koopman et al., 2008; Broeze et al., 2011; Broeze et al., 2011).

Second, we estimated the baseline prevalence of tubal pathology, diagnosed on HSG and laparoscopy, within the original individual studies before imputation, as well as for the complete imputed dataset.

Third, fecundity rate ratios (FRR) and 95% confidence intervals (95% CI) for the occurrence of (spontaneous) pregnancy for the following patient characteristics were estimated by Cox regression modelling (Cox et al., 1972): female age, duration of subfertility, body mass index (BMI), subfertility type (i.e. primary versus secondary subfertility), previous PID, previous Chlamydia infection, previous pelvic surgery or positive Chlamydia antibody test (CAT) result. Also, FRRs for the occurrence of (spontaneous) pregnancy after positive HSG or laparoscopy were estimated. A FRR expresses the probability of spontaneous pregnancy per time unit for women with an abnormal test result, relative to that probability for similar women with a normal test result. Before estimating FRRs, the proportional hazards assumption of the Cox model was checked by plotting the log minus log function and checking whether the curves were parallel. For continuous variables the assumption of a linear association with the hazard was checked, and followed by an appropriate transformation if required. In multivariable Cox regression, adjusted FRRs for the occurrence of (spontaneous) pregnancy after HSG or laparoscopy were estimated, taking into account the patient characteristics from clinical history and physical examination.

Fourth, to assess the added value of tubal patency tests in the prediction of spontaneous pregnancy, relative to patient characteristics, we built multivariable Cox regression models including female age, duration of subfertility, type of subfertility, previous PID, previous Chlamydia infection, previous pelvic surgery and a positive CAT test. Variables with a p-value below 0.30 in a univariable analysis were considered as candidate predictors in the multivariable Cox regression analysis. We used a significance level of 0.20 to keep predictors in the model. Four prediction models were created. The first model was based on the patient characteristics only, in the second and third model the patient characteristics model was extended with findings from HSG or laparoscopy, respectively while in the fourth model both tubal patency tests were added to the patient characteristics model. By comparing these models, the added value of HSG and/or laparoscopy to patient characteristics in the prediction of spontaneous pregnancy was estimated. We calculated the area under receiver operating characteristics (ROC) curves based on the estimated probabilities of pregnancy from the different models.

Except for the multivariable models, all analyses were performed for HSG and laparoscopy separately. All analyses were stratified for study center to account for the heterogeneity in
pregnancy chances across studies. P-values below 0.05 were considered to indicate statistical significance in all models.

Finally, to estimate the capacity of HSG and laparoscopy to predict spontaneous pregnancy in women with different patient profiles, we performed Cox regression analyses for spontaneous pregnancy in which interaction terms between tubal status on HSG or laparoscopy and the specific patient characteristic were added to a model containing both main effects, using the prognostic factor as a part of the interaction term. The p-value of the interaction coefficient indicates a difference in predictive capacity of the test of interest for

**Figure I.** Flowchart of included studies
different profiles of that specific patient characteristic. Data were analysed using SPSS 17.0 (SPSS Inc., Chicago, Il, USA).

Results

Literature search and data acquisition
We detected 680 potentially relevant titles on Medline, 252 on Embase and 83 in the Cochrane library. After reading the abstracts of these studies, 36 studies were deemed potentially eligible and hardcopies were requested for full reading. Of these 36 studies, 10 studies contained no diagnostic data on HSG, HyCoSy, THL or laparoscopy and two of them were systematic reviews. Six studies that included fallopscopy, salpingoscopy or radionuclide HSG had to be excluded. Five studies reported no pregnancy data and were also excluded. For three studies the articles were based on a duplicate dataset.
Ten studies from nine authors were found to meet our inclusion criteria, and these authors were approached to participate in this IPD meta-analyses by sharing their original data. In the end, eight authors could be contacted and seven of them were able to provide data, of which one author provided data from two studies (Mol et al., 1999; Mol et al., 1997; Van Tetering et al., 2007; Lindborg et al., 2009; Marianowski et al., 2007; Perquin et al., 2006; Perquin et al., 2007; van der Steeg et al., 2007). A flow chart of the inclusion of studies is shown in figure I.

Data on 6,326 individual patients from seven prospective cohort studies and one randomized controlled trial could be combined in the summary database. In all studies, included patients were referred to a fertility clinic after at least one year of unfulfilled child wish. One study

Figure II. Overview of methodological quality of reporting of included studies, according to the adjusted QUADAS checklist

134
reported on HyCoSy only, one study on THL only and six studies described both HSG and laparoscopy. For all tests, no distinction could be made between proximal and distal tubal pathology.

**Quality assessment**

The methodological quality of reporting of the studies included in the analyses, as assessed with the adjusted QUADAS checklist, is summarized in figure II. In all studies selection criteria and study design were described clearly. Also, in all studies except one, loss to follow-up was explained and duration of follow-up was reported in all studies. Three of the included studies did not describe the exclusion criteria in sufficient detail.

**Table I. Study characteristics of included studies**

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Test</th>
<th>Outcome definition</th>
<th>Study design</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mol</td>
<td>258</td>
<td>HSG</td>
<td>SP and live birth</td>
<td>Prospective cohort study</td>
<td>Newly referred subfertile women</td>
<td>Abnormal semen analysis</td>
</tr>
<tr>
<td>Van der Steeg/ Steures</td>
<td>3506</td>
<td>HSG and DLS</td>
<td>SP and live birth</td>
<td>Prospective cohort study</td>
<td>Newly referred subfertile women</td>
<td>None</td>
</tr>
<tr>
<td>Van Tetering</td>
<td>244</td>
<td>THL</td>
<td>SP</td>
<td>Prospective cohort study</td>
<td>Newly referred subfertile women</td>
<td>None</td>
</tr>
<tr>
<td>Collins</td>
<td>1495</td>
<td>HSG and DLS</td>
<td>SP</td>
<td>Prospective cohort study</td>
<td>Newly referred subfertile women</td>
<td>None</td>
</tr>
<tr>
<td>Lindborg</td>
<td>334</td>
<td>HyCoSy</td>
<td>SP and live birth</td>
<td>Prospective cohort study</td>
<td>Newly referred subfertile women</td>
<td>Previous tubal testing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Female age &gt; 40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Abnormal semen analysis</td>
</tr>
<tr>
<td>Marianowksi</td>
<td>42</td>
<td>HSG and DLS</td>
<td>SP</td>
<td>Prospective cohort study</td>
<td>Newly referred subfertile women</td>
<td>None</td>
</tr>
<tr>
<td>Perquin 2006</td>
<td>159</td>
<td>HSG and DLS</td>
<td>SP</td>
<td>Randomized controlled trial</td>
<td>Newly referred subfertile women</td>
<td>Previous tubal testing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Female age &gt; 37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Abnormal semen analysis</td>
</tr>
<tr>
<td>Perquin 2007</td>
<td>288</td>
<td>HSG and DLS</td>
<td>SP</td>
<td>Prospective cohort study</td>
<td>Newly referred subfertile women</td>
<td>Previous tubal testing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Female age &gt; 37</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Abnormal semen analysis</td>
</tr>
<tr>
<td>Total</td>
<td>6,326</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SP: spontaneous pregnancy
Table II. Baseline characteristics of included studies

Table IIA. Patient characteristics of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Age (mean)</th>
<th>Duration (median)</th>
<th>BMI (median)</th>
<th>Primary subfertility (%)</th>
<th>Previous EUG (%)</th>
<th>Previous Chlamydia (%)</th>
<th>Previous PID (%)</th>
<th>Previous pelvic surgery (%)</th>
<th>Previous abdominal surgery (%)</th>
<th>CAT positivity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mol</td>
<td>30</td>
<td>2.4</td>
<td>na</td>
<td>57</td>
<td>4.0</td>
<td>14</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>30</td>
</tr>
<tr>
<td>Van der Steeg/Steures</td>
<td>32</td>
<td>1.6</td>
<td>23</td>
<td>64</td>
<td>3.0</td>
<td>6.2</td>
<td>3.2</td>
<td>1.4</td>
<td>19</td>
<td>30</td>
</tr>
<tr>
<td>Van Tetering</td>
<td>32</td>
<td>1.4</td>
<td>23</td>
<td>74</td>
<td>na</td>
<td>3.0</td>
<td>2.8</td>
<td>0.9</td>
<td>13</td>
<td>6.8</td>
</tr>
<tr>
<td>Collins</td>
<td>30</td>
<td>3.0</td>
<td>na</td>
<td>77</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>15</td>
<td>na</td>
</tr>
<tr>
<td>Lindborg</td>
<td>32</td>
<td>2.0</td>
<td>22</td>
<td>65</td>
<td>na</td>
<td>14</td>
<td>4.5</td>
<td>1.8</td>
<td>16</td>
<td>na</td>
</tr>
<tr>
<td>Marianowski 2007</td>
<td>32</td>
<td>2.0</td>
<td>na</td>
<td>88</td>
<td>na</td>
<td>na</td>
<td>21</td>
<td>na</td>
<td>29</td>
<td>na</td>
</tr>
<tr>
<td>Perquin 2006</td>
<td>31</td>
<td>2.0</td>
<td>na</td>
<td>na</td>
<td>1.3</td>
<td>4.5</td>
<td>2.5</td>
<td>1.3</td>
<td>na</td>
<td>29</td>
</tr>
<tr>
<td>Perquin 2007</td>
<td>30</td>
<td>2.0</td>
<td>na</td>
<td>na</td>
<td>0.7</td>
<td>4.3</td>
<td>3.6</td>
<td>1.0</td>
<td>na</td>
<td>33</td>
</tr>
<tr>
<td>Overall</td>
<td>32 (18 - 46)</td>
<td>1.9 (1 - 16)</td>
<td>23 (16 - 48)</td>
<td>68</td>
<td>2.1</td>
<td>7.2</td>
<td>3.6</td>
<td>1.4</td>
<td>17</td>
<td>29</td>
</tr>
</tbody>
</table>
The analyses of the consistency between the received data and the published results showed only minimal differences in patient characteristics, which were therefore ignored. Study characteristics of the included articles are listed in table I.

### Statistical analyses
Baseline characteristics of the included women are shown in table IIA. Of the 6,326 women, 2,763 underwent HSG, 1,456 underwent laparoscopy and 1,843 underwent both tests. At HSG, the prevalence of unilateral tubal pathology was 13%, the prevalence of bilateral tubal pathology was 11% (table IIB). At laparoscopy, the prevalence of unilateral tubal pathology was 13% and the prevalence of bilateral tubal pathology at laparoscopy was 13% (table IIB). There was only fair agreement between HSG and laparoscopy (kappa 0.34) (Landis and Koch, 1977). The Cox analyses of the patient characteristics showed that female age, duration and type of subfertility were significant predictors for the occurrence of natural conception (table III). Higher female age, a longer duration of subfertility and primary subfertility were associated with a significant lower probability of natural conception. A history of PID, previous Chlamydia infection, previous pelvic surgery and a positive CAT result were no significant predictors of natural conception. For women that underwent laparoscopy, higher BMI also showed a significant lower probability of natural conception, whereas for women that underwent HSG, BMI was no significant predictor of natural conception.

The unadjusted FRR for unilateral tubal pathology at HSG was 0.83 (95% CI 0.67 – 1.05) and 0.40 (95% CI 0.28 – 0.57) for bilateral tubal pathology (table IIIA). The adjusted, multivariable FRR for HSG, which expresses the effect of HSG on top of the prognostic patient characteristics, was 0.74 (95% CI 0.50 – 1.09) for unilateral tubal pathology and 0.19 (95% CI 0.06 – 0.60) for bilateral tubal pathology.
### Table III. Results of uni- and multivariable Cox regression analysis

#### Table IIIA. HSG

<table>
<thead>
<tr>
<th>N (%)</th>
<th>Univariable (unadjusted) analysis</th>
<th>Multivariable (adjusted) analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FRR</td>
<td>95% CI</td>
</tr>
<tr>
<td>No tubal pathology</td>
<td>3292</td>
<td>1</td>
</tr>
<tr>
<td>Unilateral tubal pathology</td>
<td>558</td>
<td>0.83</td>
</tr>
<tr>
<td>Bilateral tubal pathology</td>
<td>444</td>
<td>0.40</td>
</tr>
<tr>
<td>Age</td>
<td>4302</td>
<td>0.96</td>
</tr>
<tr>
<td>Duration subfertility</td>
<td>4297</td>
<td>0.73</td>
</tr>
<tr>
<td>BMI</td>
<td>1666</td>
<td>1.00</td>
</tr>
<tr>
<td>Secondary subfertility</td>
<td>4229</td>
<td>1.34</td>
</tr>
<tr>
<td>Hx PID</td>
<td>4011</td>
<td>1.07</td>
</tr>
<tr>
<td>Hx Chlamydia</td>
<td>4008</td>
<td>0.79</td>
</tr>
<tr>
<td>Hx pelvic surgery</td>
<td>4019</td>
<td>1.25</td>
</tr>
<tr>
<td>CAT positive</td>
<td>3946</td>
<td>0.84</td>
</tr>
</tbody>
</table>

#### Table IIIB. DLS

<table>
<thead>
<tr>
<th>N (%)</th>
<th>Univariable (unadjusted) analysis</th>
<th>Multivariable (adjusted) analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FRR</td>
<td>95% CI</td>
</tr>
<tr>
<td>No tubal pathology</td>
<td>2070</td>
<td>1</td>
</tr>
<tr>
<td>Unilateral tubal pathology</td>
<td>444</td>
<td>0.47</td>
</tr>
<tr>
<td>Bilateral tubal pathology</td>
<td>464</td>
<td>0.97</td>
</tr>
<tr>
<td>Age</td>
<td>2969</td>
<td>0.77</td>
</tr>
<tr>
<td>Duration subfertility</td>
<td>2967</td>
<td>0.96</td>
</tr>
<tr>
<td>BMI</td>
<td>879</td>
<td>1.35</td>
</tr>
<tr>
<td>Secondary subfertility</td>
<td>2905</td>
<td>1.31</td>
</tr>
<tr>
<td>Hx PID</td>
<td>2740</td>
<td>0.89</td>
</tr>
<tr>
<td>Hx Chlamydia</td>
<td>2739</td>
<td>1.13</td>
</tr>
<tr>
<td>Hx pelvic surgery</td>
<td>2742</td>
<td>0.78</td>
</tr>
</tbody>
</table>

The unadjusted FRR for unilateral tubal pathology at laparoscopy was 0.84 (95% CI 0.61 – 1.16), and 0.47 (95% CI 0.32 – 0.70) for bilateral tubal pathology (table IIIB). The adjusted, multivariable FRR for laparoscopy, in which the effect of laparoscopy was corrected for
patient characteristics, was 0.77 (95% CI 0.55 – 1.06) for unilateral tubal pathology and 0.45 (95% CI 0.30– 0.67) for bilateral tubal pathology.

To assess the added value of tubal patency tests in the prediction of spontaneous pregnancy, relative to patient characteristics, we built four multivariable Cox regression models. The multivariable model based on patient characteristics, female age, duration of subfertility, BMI, type of subfertility, previous Chlamydia infection and a positive CAT result had an area under the ROC curve of 0.61 (95% CI 0.60 - 0.63). Adding HSG and/or laparoscopy results to the patient characteristics model increased the predictive capacity of this model (AUCs of 0.63 (95% CI 0.62 – 0.65), 0.62 (95% CI 0.60 – 0.63) and 0.63 (95% CI 0.62 - 0.65) for adding HSG, laparoscopy and both HSG and laparoscopy, respectively).

In a final step we evaluated whether the prognostic effect of tubal pathology at HSG or laparoscopy varied for women with different profiles. For both unilateral and for bilateral tubal pathology at HSG we observed significant interactions with female age (p = 0.03), previous PID (p = 0.05), previous Chlamydia infection (p = 0.02) and a positive CAT result (p < 0.01). For female age and duration of subfertility the changes the FRRs of HSG for women with different profiles are shown in figure IIIA.

Figure IIIB shows the changes in the FRRs of laparoscopy for women with different profiles. The addition of interaction terms to the Cox regression for pregnancy after laparoscopy showed significant interactions for female age (p = 0.02), BMI (p = 0.01), previous Chlamydia infection (p < 0.01) and positive CAT results (p = 0.03).

For women with unilateral tubal pathology at HSG or laparoscopy, increasing age leads to lower rates of natural conception. For women with bilateral tubal pathology natural conception rates were low for all ages. Longer duration of subfertility lead to lower rates of natural conception for women with both unilateral or bilateral tubal pathology at HSG or laparoscopy.
Figure III. Prediction of pregnancy at different profiles
Figure IIIA. Prediction of pregnancy after HSG.

**FRR female age (HSG)**

![Graph showing FRR for female age with different tubal pathologies](image)

FRR of HSG for women with different female age.

**FRR duration subfertility (HSG)**

![Graph showing FRR for duration of subfertility with different tubal pathologies](image)

FRR of HSG for women with different duration of subfertility.
Figure IIIb. Prediction of pregnancy after DLS

FRR female age (DLS)

FRR of laparoscopy for women with different female age.

FRR duration subfertility (DLS)

FRR of laparoscopy for women with different duration of subfertility.
Discussion

To provide optimal management for couples with fertility problems, it is essential to know the chances of natural conception, since the natural conception rate guides decisions about assisted reproductive therapy. So far, the added value of tubal patency tests to information from clinical history and physical examination in the prediction of natural conception was relatively unknown. Also, the predictive value of these tests has never been investigated for its consistency between women with different profiles.

The current prognostic IPD meta-analysis shows that the integration of the results of HSG with the available patient characteristics shows the best prediction of natural conception. Addition of laparoscopy does not further improve the predictive value of the model. The capacity of HSG and laparoscopy to predict natural conception is influenced by female age, BMI, previous PID, previous Chlamydia infection and a positive CAT. In women with unilateral tubal pathology at HSG or laparoscopy, increasing age lowers their natural conception rate, while in women with bilateral tubal pathology, increasing female age is a less important factor in predicting natural conception. Longer duration of subfertility leads to lower rates of natural conception for all women.

The data of the current study open up the possibility to combine data from clinical history and physical examination with tubal patency test results to estimate the best prediction of natural conception after the diagnostic work-up in subfertile couples. This assessment reflects daily practice in which the diagnostic work-up is a consecutive process. The use of data from multiple studies performed in different countries allows a general applicable estimation of natural conception rates after tubal patency tests. Also, for continuous variables such as female age and duration of subfertility, the complete range of values could be used in the analyses, showing the chances of natural conception for every possible age and duration of subfertility.

A limitation of IPD meta-analyses is their dependency of the availability of original data. As shown in the flowchart, not all authors that were approached were able to provide their data for this meta-analysis and therefore the prognostic models created in this study could only be based on the studies made available. Although the final summary database reflects daily practice, where not all centers perform the same diagnostic tests or note the same patient information in their charts, the lack of more data might hamper the analyses. From a previous study, for example, it is known that endometriosis and previous ectopic pregnancy are strong risk indicators for tubal pathology (Coppus et al., 2007). Unfortunately, most databases did not contain data on endometriosis or ectopic pregnancy, thus limiting the possibility to assess these issues in our analyses.

To be able to use as much data as possible, a multiple imputation approach was chosen for variables with less than 50% missing values, which improves the model analyses. In this way missing patient characteristics and missing test results were imputed, since omitting a predictor (i.e. patient characteristic) with missing values from a multivariable analysis
decreases its ability and imputation is better than a complete case analysis (Janssen et al., 2010; Moons et al., 2006). Also, due to a lack of long term follow-up data, we decided to analyse pregnancy rates up to 12 months after the performance of HSG or laparoscopy. Less data were available for follow-up longer than one year and therefore, analyses would otherwise be hampered by small amounts of patients left at later time points. We also described the issue of missing data in IPD meta-analyses in our previous diagnostic meta-analyses (Broeze et al., 2011; Broeze et al., 2011; Broeze et al., 2012). Hopefully, better models may be developed in the future, once more variables could be provided by authors of the original studies. We therefore stress the importance of better data sharing policies as well as advancing data storage facilities.

From a clinical perspective, the definition of tubal pathology influences the conception rate as predicted based on the tubal patency tests. Our definition of tubal pathology included occlusion of the fallopian tubes, with or without hydrosalpinges and peritubal adhesions. We did not discriminate between proximal and distal tubal occlusion, since clinical management and pregnancy chances are similar in both types of occlusion (Farhi et al., 2007). Also, it is important to realize that several Chlamydia antibody tests exists, of which ELISA, MIF and IF were performed in the studies included in this meta-analysis. These tests all have different accuracies, indicating that different CAT tests might have different abilities to predict tubal pathology. Tests that are specific for Chlamydia trachomatis show other accuracies than tests that show cross-reactivity with Chlamydia pneumoniae, leading to higher false positive rates (Land et al., 2003). Most women in this study underwent ELISA manufactured by Medac, which has the least cross-reactivity with Chlamydia pneumoniae. In a previous study it was shown that subfertile women with a positive CAT and without tubal pathology at HSG or laparoscopy have a 33% lower probability of spontaneous pregnancy as compared to women with a negative CAT (Coppus et al., 2011). In the present meta-analysis CAT alone was not a significant predictor of natural conception, but a positive CAT influenced the predictive value of both HSG and laparoscopy, resulting in lower natural conception rates. Since unilateral tubal pathology does not seem to predict natural conception rates, tubal patency tests should be performed to detect women with bilateral tubal pathology, so that they can be treated with IVF to increase their pregnancy chances. Furthermore, the prediction of natural conception by tubal patency tests in this meta-analysis was similar for HSG and laparoscopy and no additional predictive value was seen when laparoscopy was performed after HSG. This corresponds to the results of Lavy et al, where in 95% of patients, laparoscopic findings did not change the HSG-based treatment plan (Lavy et al., 2004). Obviously, laparoscopy could still have additional value in women with endometriosis or in women with inconclusive results at HSG.

In conclusion, the present IPD meta-analysis shows that natural conception can be predicted from patient characteristics such as female age, duration and type of subfertility, BMI, previous Chlamydia infection and a positive CAT result. Addition of HSG to these patient characteristics slightly increased the predictive capacity, whereas laparoscopy had no added
value to HSG in these women. Increasing age leads to lower rates of natural conception for women with unilateral tubal pathology, whereas for women with bilateral tubal pathology, natural conception rates were low for all ages. Longer duration of subfertility also leads to lower rates of natural conception for women with unilateral or bilateral tubal pathology. Since no added value of laparoscopy was found once HSG was performed after notification of patient characteristics, we conclude that patient characteristics combined with HSG results has the best predictive capacity for natural conception.
References


Moons KGM, Donders RART, Stijnen T, and Harrell FE jr. Using the outcome for imputation of missing predictor values was preferred. J Clin Epi 2006;59:1092-1101.


