Diagnostic tests for tubal pathology from a clinical and economic perspective
Verhoeve, H.R.

Citation for published version (APA):
Verhoeve, H. R. (2013). Diagnostic tests for tubal pathology from a clinical and economic perspective

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: http://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.
Cost-effectiveness of tubal patency tests

Harold Verhoeve
Lobke Moolenaar*
Peter Hompes
Fulco van der Veen
Ben Willem Mol

BJOG 2013, in press

* Shared first authorship
Abstract

Background
Guidelines are not in agreement on the most effective diagnostic scenario for tubal patency testing. Therefore we evaluated the cost-effectiveness of invasive tubal testing in subfertile couples compared with no testing and treatment.

Methods
We performed a cost-effectiveness analysis, using a decision analytic framework based on a computer-simulated cohort of subfertile women. We evaluated six diagnostic and treatment scenarios: [1] No tests and no treatment [2] Immediate treatment without tubal testing [3] Delayed treatment without tubal testing [4] Hysterosalpingography (HSG), followed by immediate or delayed treatment, according to diagnosis (tailored treatment) [5] HSG and a diagnostic laparoscopy (DL) in case HSG does not prove tubal patency, followed by tailored treatment [6] DL followed by tailored treatment. Main outcome was the expected cumulative live births after three years. Secondary outcomes were costs per couple and the incremental cost-effectiveness ratio.

Results
For a 30 year old woman, with otherwise unexplained subfertility for 12 months, three year cumulative live birth rates were 51.8%, 78.1%, 78.4%, 78.4%, 78.6% and 78.4% and costs per couple were €0, € 6,968, € 5,063, € 5,410, € 5,405 and € 6,163 for scenario 1, 2, 3, 4, 5 and 6, respectively. The incremental cost-effectiveness ratio compared with scenario 1 (reference strategy), were € 26,541, € 19,046, € 20,372, € 20,150 and € 23,184 for scenario 2, 3, 4, 5 and 6, respectively. Sensitivity analysis showed the model to be robust over a wide range of values for the variables.

Conclusions
The most cost-effective scenario is to perform no diagnostic tubal tests and to delay IVF treatment for at least 12 months for women aged until 38 years and to perform no tubal tests and start immediate IVF treatment from the age of 39 years. If an invasive diagnostic test is planned, HSG followed by tailored treatment or a diagnostic laparoscopy in case HSG shows no tubal patency, is more cost-effective than diagnostic laparoscopy.
Introduction

In couples with subfertility, defined as one year of regular unprotected intercourse without conception, guidelines recommend the instigation of basic diagnostic investigations for both women and men (ASRM Practice Committee Reports, 2006; National Institute for Clinical Excellence, 2004). In most guidelines the last step in this fertility work-up consists of assessment of tubal patency.

The guidelines are not in agreement on the most effective diagnostic scenario for tubal patency testing, i.e. there is no consensus on which test should initially be used, and there is no consensus on the sequence of tests in the fertility work-up (ASRM Practice Committee Reports, 2006; Bosteels et al., 2007; den Hartog et al., 2008; Fatum et al., 2002; National Institute for Clinical Excellence, 2004; Perquin et al., 2006). The reason for this is that all diagnostic tests for the detection of tubal pathology have limitations. Like many diagnostic test studies, studies on tubal patency tests are hampered by verification and review bias. Also, variations in demographic features, disease prevalence, observer and instrument variation, make interpretation of the tests difficult (Lijmer et al., 1999; Whiting et al., 2004). Finally, the guidelines do not consider costs or the combination of cost and effectiveness.

This is worrisome, because the decision whether to start treatment or not is highly dependent on tubal patency. Subfertile women who are diagnosed with bilateral tubal occlusion are usually advised IVF treatment, where women with patent tubes are advised to adhere to a period of expectant management if they have a favourable prognosis to accomplish a natural conception, or are advised to start with superovulation and insemination when the natural conception chances are less favourable (Hunault et al., 2004; Steures et al., 2006).

In view of this lack of agreement the aim of our study was to compare cost and effectiveness of different invasive diagnostic scenarios for the detection of tubal pathology. For this purpose we developed a Markov analytic model (Briggs et al., 2006).

Materials and methods

We constructed a Markov decision tree for subfertile couples who finished their basic fertility work-up except for tubal assessment. A Markov model is a more complicated decision model used to analyze recurring events over time. It can be very useful for the evaluation of cost-effectiveness analyses in reproductive medicine, because every month there is a new chance to conceive. Markov models can be used to compute the costs per live birth and the incremental cost-effectiveness ratio (ICER). The incremental cost-effectiveness ratio represents the extra costs per live birth between two scenarios. These costs are calculated by dividing the differences in costs by the differences in live births of two scenarios. Normal practice is to order strategies
or scenarios from the least to the most effective. Dominated strategies are then eliminated and the ICERS are then calculated for each strategy in comparison with its next best alternative.

**Patient characteristics and diagnostic tests**

Our base-case calculation was centred on a 30 year old women with a regular menstrual cycle, who had finished the initial fertility work-up except for the assessment of tubal patency and had a partner with a normal semen analysis. This initial work-up included a medical history, physical examination, evaluation of the menstrual cycle and a semen analysis.

Thereafter a choice had to be made whether to offer a hysterosalpingogram (HSG) or a diagnostic laparoscopy (DL) to assess the tubal status of the woman, or not to offer a diagnostic test. HSG and DL are both invasive tests. At HSG a contrast medium is slowly injected through the cervical canal into the uterine cavity. By X-ray imaging the flow of contrast can be followed and the uterine cavity and lumen of the Fallopian tubes can be visualized. For diagnostic laparoscopy two abdominal entrance-ports are required to introduce the laparoscope and an instrument. A water-soluble dye is then injected into the uterine cavity and tubal patency can subsequently be assessed through the laparoscope. Also, the pelvic cavity can be inspected for presence of adhesions and endometriosis. HSG is typically performed in an outpatient setting and can be painful. DL requires general anaesthesia, theatre time and more personnel.

**Diagnostic scenarios**

Because there is a wide variation in clinical practice in the timing of tubal patency tests and in the type of test that is preferred by the provider, we first analysed the implications of these choices on pregnancy outcome and costs. To do so we defined six possible clinical scenarios between which a clinician can choose (Figure 1). Although the scenario of expectant management and no diagnostic test is not usually chosen, we used this scenario as the reference to enhance potential differences between all scenarios. To assess the effect of patient characteristics on these choices, we performed sensitivity analyses, which will be described later.

Scenario 1 was the reference scenario. In this scenario no tubal assessment was performed and no treatment (i.e. IVF) was offered. Couples had a probability to conceive naturally for three years.

Scenario 2 represented a scenario in which no diagnostic test for tubal pathology was performed and all women received immediate treatment to a maximum of 3 IVF treatments. In case no live birth was achieved, couples still had a probability to conceive naturally.

Scenario 3 was defined as a delayed treatment scenario. In this scenario no diagnostic test for tubal pathology was offered. All women were subjected to expectant management for one year. If no live birth was accomplished in this period of time, IVF treatment was started to a
**Figure 1**
Overview of the model

### Scenario 1
No diagnostics & no treatment
- **Age 30 yrs**
- Expectant management (36 cycles)

### Scenario 2
No diagnostics & immediate treatment
- Age 33 yrs
- Expectant management (36 cycles)

### Scenario 3
No diagnostics & delayed treatment
- **Age 30 yrs**
  - Expectant management (36 cycles)

### Scenario 4
HSG & tailored treatment
- **Age 33 yrs**
  - Expectant management (36 cycles)

### Scenario 5
HSG and DL for no tubal patency at HSG
- Expectant management (36 cycles)

### Scenario 6
DL & tailored treatment
- No patent tube
- Expectant management (36 cycles)

**COST-EFFECTIVENESS ANALYSIS**
maximum of 3 cycles. After failed IVF treatment couples were still considered to have a probability to conceive naturally. In scenario 4, all women received an HSG to assess their tubal status. According to diagnosis, a treatment scenario of delayed or immediate treatment was followed (tailored treatment). Delayed treatment consisted of one year of expectant management followed by a maximum of three IVF treatments. Immediate treatment with a maximum of 3 IVF treatments was followed by expectant management in case of IVF failure. Women with bilateral tubal pathology on HSG received immediate treatment and women with unilateral tubal pathology or no tubal pathology on HSG were subjected to the delayed treatment. In Scenario 5 all women received a hysterosalpingogram and those women with bilateral tubal pathology were subjected to a diagnostic laparoscopy. In this scenario laparoscopy followed as soon as possible after the HSG. Depending on diagnosis tailored treatment was assigned. Scenario 6 represented a scenario in which all women underwent a diagnostic laparoscopy and tailored treatment was assigned according to diagnosis. For all six scenarios, the costs of diagnostic tests, treatment and pregnancy probabilities were calculated over a three year period after finishing the initial fertility work-up. Cycle length was set at one month. We used a time frame of 3 years, because in couples with unexplained subfertility natural conception chances within these 3 years are substantial (Pandian et al., 2005). The outcome was defined as live birth of at least one child. Estimates for the base-case scenario and ranges for sensitivity analyses are summarized in Table 1 and were derived from peer reviewed literature, as referenced. Rates were converted to probabilities when needed. The cost calculation was made according to the Dutch situation in the year 2009; hence costs were adjusted according to the consumer price index (Statistics the Netherlands, 2011). We assumed that in this period no significant cost changes in the treatment protocol occurred except for inflation. The model was built from a health care perspective.

**Details of computer simulation model**

**Tubal pathology**

In the model we differentiated between women with bilateral or no bilateral tubal pathology (including unilateral pathology). This distinction was made because women with bilateral tubal pathology have a severe reduction in natural pregnancy chances in contrast to unilateral tubal pathology (Mol et al., 1999; Verhoeve et al., 2011). DL was used as the reference test. Sensitivity and specificity of HSG were used to predict the possibility of true positive and false positive findings in women after HSG. The sensitivity and specificity of HSG and the prevalence of bilateral tubal pathology in the subfertile population were extracted from an individual patient data meta-analysis (Table 1) (Broeze et al., 2012). For the calculation of our main outcome we assumed that the prevalence of bilateral tubal pathology was 12% (Broeze et al., 2012). We also assumed that no complications occurred with either HSG or DL.
Live Birth probabilities
We derived live birth probabilities for IVF from a prospective cohort study performed between 2002 and 2004 (Lintsen et al., 2007; Moolenaar et al., 2011). In our analysis we considered up to three treatments because this is reimbursed by Dutch Healthcare services and is common practice in Europe (Andersen et al., 2007). Live birth rates after natural conception were calculated with the Hunault prediction model which takes female age, primary or secondary subfertility, duration of subfertility, percentage of motile sperm and referral by general practitioner or gynaecologist into account (Hunault et al., 2004). The Hunault model calculates the live birth rate after 1 year of expectant management for women without tubal pathology. We made the following assumptions: women were subfertile for 1 year, were primary subfertile and their partners had a normal semen analysis; women with bilateral tubal pathology were not able to conceive naturally; and live birth probabilities declined with advancing age according to the model of Hunault. Live birth rates per year were converted to live birth rates per month.

Costs
Costs per cycle were derived from the Dutch Umbrella study on fertility treatments and from our institutional data (Merkus, 2006; The Dutch Healthcare Authority, 2009). All monetary units were converted into the equivalent of 2009 using the consumer price index (Statistics the Netherlands, 2011). We assumed that in this period no significant cost changes in the treatment protocol occurred except for inflation (Table 1).

Outcomes
Cumulative live birth rates were determined for each scenario, as were the estimated costs. Live birth rates were defined as ongoing pregnancies resulting in a live birth. There was no distinction between singleton or multiple gestations. Using these values, we computed the costs per live birth and the incremental cost-effectiveness ratio (ICER). The incremental cost-effectiveness ratio represents the extra costs per live birth between two scenarios. These costs are calculated by dividing the differences in costs by the differences in live birth rate of two scenarios.

Sensitivity analysis
To address the uncertainty regarding our assumptions we carried out one-way and probabilistic (Monte Carlo simulations) sensitivity analyses. In one-way sensitivity analysis we varied all variables independently. A threshold analysis was performed to determine if and when a variable changed the threshold value. This represents the value of a variable above which another scenario is preferred.

In our base case calculation no discounting was applied, since it would have little influence on the outcome, as most costs were made within one year. In one-way and multi-way sensitivity analysis we tested the effect of different discounting rates for costs and effects.
In our model, live birth rates were adjusted over time, in our sensitivity analyses we assumed that this was fixed during the period of the model, i.e. live birth rates of a 30 year women declined with the same percentage as live birth rates of a 31 year old woman.

In our base case calculations no considerations were made about endometriosis. For that reason we also tested the impact of finding minimal to mild endometriosis at diagnostic laparoscopy and coagulating the lesions in the same session on live birth rates. To adjust the live birth rates, we used an OR of 1.64 (CI 1.05-2.57) if women were treated for endometriosis (Jacobson et al., 2010). The probability of finding minimal or mild endometriosis at DL was set at 30% (Meuleman et al., 2009).

In probabilistic sensitivity analysis the uncertainty of each parameter is quantified in terms of a probability distribution of this parameter. Using probabilistic sensitivity analysis, confidence intervals of the outcome are computed. For this analysis distributions were fitted for all parameters in the model. We fitted beta-distributions for probabilities. If we were not able to set beta distributions, normal distributions were fitted. The normal distributions were calculated according to the confidence interval from the study or by the plausible range provided by expert opinion (Briggs et al., 2006). For the probabilistic sensitivity analysis, 5000 iterations of 5000 women were performed. To visualize the probability of the optimal scenario based on the willingness to pay (WTP), cost-effectiveness acceptability curves were computed. The willingness to pay expresses the payers willingness to pay for the outcome of concern, i.e. in this study for a live birth. In reproductive medicine the WTP for a live born is an important unknown, but in health care questions the WTP is usually expressed as one extra quality of adjusted life year gained. Because this is usually set by decision makers at € 80,000, we used this figure as set point for a live birth (Council for Public Health, 2006). The ranges and values of all variables used in the sensitivity analyses are shown in Table 1.

To address the effect of varying female age, the model was also applied to women aged 20 to 43 years old. We performed our analysis by using a computer-generated Markov model (Tree Age Pro 2009, Tree Age Inc, Williamstown, MA, USA). No approval for this research was needed.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Base case assumption</th>
<th>Distribution (*)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Natural live birth rate</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No tubal pathology</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>Probability/cycle</td>
<td></td>
<td>Reference</td>
</tr>
<tr>
<td>30</td>
<td>0.034</td>
<td>Normal (0.017-0.050)</td>
<td>Hunault <em>et al.</em>, 2004</td>
</tr>
<tr>
<td>31</td>
<td>0.027</td>
<td>Normal (0.014-0.041)</td>
<td>Hunault <em>et al.</em>, 2004</td>
</tr>
<tr>
<td>32</td>
<td>0.021</td>
<td>Normal (0.010-0.031)</td>
<td>Hunault <em>et al.</em>, 2004</td>
</tr>
<tr>
<td>33</td>
<td>0.016</td>
<td>Normal (0.008-0.024)</td>
<td>Hunault <em>et al.</em>, 2004</td>
</tr>
<tr>
<td><strong>IVF live birth rate</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cycle</td>
<td>Probability/cycle</td>
<td>Distribution (*)</td>
<td>Reference</td>
</tr>
<tr>
<td>Age 30</td>
<td>1</td>
<td>0.277</td>
<td>Normal (0.14-0.42)</td>
</tr>
<tr>
<td>2</td>
<td>0.253</td>
<td>Normal (0.13-0.38)</td>
<td>Lintsen <em>et al.</em>, 2007</td>
</tr>
<tr>
<td>3</td>
<td>0.219</td>
<td>Normal (0.11-0.33)</td>
<td>Lintsen <em>et al.</em>, 2007</td>
</tr>
<tr>
<td><strong>Prevalence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilateral tubal pathology</td>
<td></td>
<td>0.12</td>
<td>Beta (α = 586 β = 4297)</td>
</tr>
<tr>
<td><strong>Sensitivity &amp; specificity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hysterosalpingography (HSG)</td>
<td>Sensitivity</td>
<td>Probability</td>
<td>Distribution (*)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.60</td>
<td>Beta (α = 2930 β = 1953)</td>
</tr>
<tr>
<td></td>
<td>Specificity</td>
<td>0.91</td>
<td>Beta (α = 4444 β = 439)</td>
</tr>
<tr>
<td><strong>Costs</strong></td>
<td>Costs</td>
<td>Distribution (*)</td>
<td>Reference</td>
</tr>
<tr>
<td>Hysterosalpingography</td>
<td>175</td>
<td>Normal (€ 150-€ 200)</td>
<td>The Dutch HCA, 2009</td>
</tr>
<tr>
<td>Diagnostic laparoscopy</td>
<td>1100</td>
<td>Normal (€ 800-€ 1300)</td>
<td>The Dutch HCA, 2009</td>
</tr>
</tbody>
</table>

* Used plausible range in case (-) and used alpha and beta (α β) in case of a beta distribution
CHAPTER 7

Results

For a 30 year old woman cumulative live birth after three years was 51.8% for scenario 1, 78.1% for scenario 2, 78.4% for scenario 3, 78.4% for scenario 4, 78.6% for scenario 5 and 78.4% for scenario 6. Costs per couple were € 0, € 6,968, € 5,063, € 5,410, € 5,405 and € 6,163 and costs per live birth were € 0, € 8,927, € 6,459, € 6,904, € 6,874 and € 7,862, for scenario 1, 2, 3, 4, 5 and 6, respectively. The incremental cost-effectiveness (ICER), i.e. the extra cost per live birth compared with scenario 1, i.e. no treatment and no diagnostic evaluation for tubal patency – the reference scenario – were € 26,541, € 19,046, € 20,372, € 20,150 and € 23,184 for scenario 2, 3, 4, 5 and 6, respectively.

Under the baseline assumptions, the model favoured scenario 3. Scenario 5 was more expensive but also more effective compared with scenario 3. The incremental cost-effectiveness for

Figure 2
Cost-effectiveness plane

- Scenario 1 No diagnostics & no treatment
- Scenario 4 HSG & tailored treatment
- Scenario 2 No diagnostics & immediate treatment
- Scenario 3 No diagnostics & delayed treatment
- Scenario 5 HSG and DL for bilateral tubal pathology at HSG
- Scenario 6 DL & tailored treatment
- ICER
scenario 3 compared with scenario 1 was € 19,046, the ICER of scenario 5 compared with scenario 3 was € 143,448. Scenario 6 was always dominated by the other scenario’s (Figure 2).

Sensitivity analyses
Univariable sensitivity analysis showed that varying the duration of expectant management prior to the start of treatment was of little influence. Scenario 3 remained the most cost-effective, provided that expectant management was adhered to for at least 2 months. Also by changing the costs of IVF, HSG or DL, no threshold could be found where a different scenario became more cost-effective. Applying univariable sensitivity analysis for different discount rates for costs, a threshold was found at a discount rate of 39.5%. Above this threshold immediate treatment became the most cost-effective strategy (scenario 2) (meaning that if costs would increase per year by 39.5% or more, immediate treatment would become more cost-effective). If costs were discounted with 4% per year and effects with 1.5% per year, delayed treatment remained the most cost-effective strategy with an ICER of € 19,279 (scenario 3). In our model live birth rate declined per year, in our sensitivity analysis we assumed that the live birth rate for the duration of the model was fixed. If IVF live birth rates would decline with more than 69% per year, HSG and DL (scenario 5) became the most cost-effective scenario. In case the probability to conceive naturally increased or decreased, our conclusions would not change.

Women with bilateral tubal pathology at DL were assumed to have no natural conception chances in our model. This is debatable since a prospective study found a fecundity rate ratio (FRR) of 0.24 for women with bilateral tubal pathology at DL (Verhoeve et al., 2011). To evaluate this effect we addressed this in a sensitivity analysis. We found that applying a FRR with a range between 0 and 1, did not change our conclusions.

The impact of finding minimal or mild endometriosis at diagnostic laparoscopy and the positive effect on the live birth rates if this was treated, did not affect our conclusions, since delayed treatment (scenario 3) remained the most cost-effective strategy, although the ICER for scenario’s 5 and 6 dropped to € 19,921 and € 20,574.

Other sensitivity analyses also did not alter the conclusions of the model, since threshold values were not found within the set ranges.

Probabilistic sensitivity analysis
The results of the probabilistic sensitivity analysis remained stable for our model and did not alter our baseline results (Table 2). The cost-effectiveness acceptability curve showed that if the willingness to pay (expressing the payers willingness to pay for a live birth) was assumed to be less than € 19,600, scenario 1 had the highest probability to be the most cost-effective scenario. If the willingness to pay would rise above € 19,600 per extra live birth, scenario 3 had the highest probability to be cost-effective.
Table 2
Outcome of the probabilistic sensitivity analysis

<table>
<thead>
<tr>
<th>Scenario*</th>
<th>Live Birth Mean (95% CI)</th>
<th>Cost per couple Mean (95% CI)</th>
<th>ICER Mean (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>0.508 (0.310-0.655)</td>
<td>-</td>
<td>Reference scenario</td>
</tr>
<tr>
<td>No diagnostics &amp; no treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 2</td>
<td>0.772 (0.602-0.886)</td>
<td>€ 6,972 (€ 4,415-€ 9,742)</td>
<td>€ 29,632 (€ 12,445-€ 67,134) dominated</td>
</tr>
<tr>
<td>No diagnostics &amp; direct treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 3</td>
<td>0.775 (0.605-0.889)</td>
<td>€ 5,086 (€ 3,125-€ 7,360)</td>
<td>€ 20,139 (€ 9,744-€ 46,037)</td>
</tr>
<tr>
<td>No diagnostics &amp; delayed treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 4</td>
<td>0.775 (0.604-0.888)</td>
<td>€ 5,431 (€ 3,432-€ 7,728)</td>
<td>€ 24,150 (€ 10,436-€ 48,549) dominated</td>
</tr>
<tr>
<td>HSG &amp; tailored treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 5</td>
<td>0.778 (0.608-0.890)</td>
<td>€ 5,428 (€ 3,455-€ 7,712)</td>
<td>€ 22,389 (€ 10,612-€ 47,255)</td>
</tr>
<tr>
<td>HSG and DL for no tubal patency at HSG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 6</td>
<td>0.775 (0.606-0.889)</td>
<td>€ 6,187 (€ 4,172-€ 8,459)</td>
<td>€ 23,150 (€ 12,595-€ 54,005) dominated</td>
</tr>
<tr>
<td>DL &amp; tailored treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* All options referenced to a common baseline, scenario 1

Female age
The incremental cost-effectiveness per additional live birth are presented in Figure 3. Up to the woman’s age of 38 years the model favoured scenario 3 as the most cost-effective scenario. From the age of 39 years scenario 2 provided the best value for money. The willingness to pay would have to rise to a maximum of € 78,507 per additional live birth for a 43 year old woman (scenario 2). Scenario 6 showed to be less cost-effective than scenarios 4 and 5 for all ages. At the age of 30 extra costs per additional live birth are minimal between the scenarios, which is caused by the minimal decline in birth rates if women become one year older. Since the increase in cumulative birth rates flattens, due to aging of the woman, over a longer period of time the differences in costs between the scenarios with advancing age increase.
Prevalence of tubal disease

Until the age of 30, delayed treatment was the most cost-effective scenario, regardless of the prevalence of tubal pathology. From the age of 40 immediate treatment became the most cost-effective treatment, regardless of the prevalence. Between the age of 31 and 39 the prevalence threshold at which immediate treatment became cost-effective declined gradually. For example, if for a woman aged 31 the prevalence of tubal pathology rose above 95% (e.g. in case the medical history revealed bilateral salpingectomy), immediate treatment became cost-effective, whereas for a woman aged 39, immediate treatment became cost-effective if the prevalence of tubal pathology rose above 10% (Figure 4).
Discussion

Because of a lack of consistency on the reported accuracy and predictive capacity of tubal patency tests, the clinical application of these tests varies widely between assisted conception units. Because of ever increasing healthcare costs and limited resources, balanced decisions in the ordering of diagnostic tests and the provision of effective treatment are becoming more important to contain expenditure.

In this study we structured the evidence on clinical and economic outcomes with a decision model technique to assess the cost-effectiveness of various scenarios for the evaluation of tubal patency. This may help in the clinical decision process as to which tubal tests should be used in the fertility work-up and when they should be used.

Our Markov model showed that the most cost-effective scenario is, to perform no diagnostic tubal tests and delay treatment for women aged until 38 and perform no diagnostic tests and start immediate treatment from the age of 39 years. The second best alternative was HSG followed by DL if necessary (scenario 5) with an ICER of € 143,448. We assumed the willingness to pay (WTP) for an extra live birth was € 80,000 (which is the accepted WTP for one extra
quality of adjusted life year in cost-effectiveness analyses (Council for Public Health, 2006). However, it is unknown what the payer (provider or couple) is prepared to pay for an extra live birth. Scenario 5 could be the most cost-effective alternative if the WTP would amount to €143,448 per extra live birth.

Until the age of 30 years and from the age of 40 the prevalence of bilateral tubal pathology was of no influence on the outcome. In-Between, immediate treatment and no diagnostic tests would be more cost-effective than delayed treatment if the prevalence of tubal pathology would be above 95% at the age of 31 years with a gradual decline of this threshold value to 10% at the age of 39 years. Similarly, the outcome was not influenced by the presence and treatment of minimal or mild endometriosis. If tubal tests would be performed, HSG and DL for those women with bilateral tubal pathology at HSG, followed by tailored treatment, was more cost-effective than diagnostic laparoscopy followed by tailored treatment. The scenarios in which only HSG (scenario 4) or DL (scenario 6) were performed, were always less cost-effective than the scenarios in which an HSG was followed by DL (in case of abnormal findings at HSG).

A strength of this study is that we used the prognostic model of Hunault for natural conception in subfertile couples (Hunault et al., 2004). This model has been shown reliable through external validation in a large subfertile population (van der Steeg et al., 2007). Also, we used the results of a prospective study on the prediction of IVF and ICSI on pregnancy outcome in women with an indication for this treatment in accordance with the national Dutch guidelines (Lintsen et al., 2007). Another strength is that we used a recent individual patient data (IPD) meta-analysis in which the information of the test characteristics of HSG was available (Broeze et al., 2012). The outcome of the various strategies were therefore based on observational studies of different, but comparable, populations. In the only other known study on the cost-effectiveness of diagnostic tests for tubal patency, which concluded that both HSG and/or laparoscopy are cost-effective in most couples, patient characteristics were not taken into account and the prognostic profiles of the couples were not externally validated. Also, the population studied, consisted of a tertiary hospital population only (Mol et al., 2001).

Some other issues also need to be addressed. First, we did not consider the Chlamydia Antibody Test (CAT). Although CAT can be of help to differentiate between women who have a low or high risk of tubal pathology, the test result is not conclusive whether tubal pathology is present and does not provide additional anatomical information like HSG does. In our sensitivity analysis we tested the effect of having a low or high probability of tubal pathology, reflected by the prevalence. Another reason not to consider CAT as a separate scenario is because a recent IPD-analysis showed that the combination of CAT and HSG is a better predictor of bilateral tubal pathology than CAT (Broeze et al., 2012).

Second, we assumed that women underwent three consecutive IVF/ICSI treatment cycles without delay in between cycles and we did not take into account frozen embryo transfers and treatment dropouts. However, by making use of the results of a large cohort of women, treated
in the Netherlands according to our national guideline for the indications of IVF/ICSI, the prognostic values are representative for the clinical practice in our country (Lintsen et al., 2007).

Third, we did not take into account the extra costs that can be involved with multiple pregnancies. We did not do so, because the current trend is to try to reduce multiple pregnancies by applying elective single embryo transfers (eSET), depending on the quality of the embryo and on female age. In the Netherlands this has resulted in a significant reduction of multiple pregnancies without a decline in pregnancy rate (Nederlandse Vereniging voor Obstetrie en Gynaecologie, 2011). The success rates of eSET are partly determined by more transfer procedures (fresh and frozen). eSET thus incurs additional costs, because these cycles require monitoring of the cycle. However, evidence exists that the costs of eSET are lower than those of double embryo transfers (DET) in the short term because of an increase in multiple pregnancies in DET (Scotland et al., 2011; Veleza et al., 2009). Research is ongoing pertaining the long term cost-effectiveness involved with eSET versus DET (van Heesch et al., 2010).

Fourth, we used the healthcare costs that are paid for by the health insurances in the Netherlands. Rates are fixed and women are reimbursed. In countries or states where there are no national health insurances, fees for consultation, diagnostic tests and fertility treatment might differ and therefore influence the costs as presented in this study. Therefore we presented our results in a disaggregated way so that individual decision makers can interpret the costs and benefits from their particular viewpoints (Drummond and Jefferson, 1996). We assume that relative differences between costs are similar and that therefore our conclusions are also applicable to societies with private healthcare systems or when fertility tests and treatment have to be funded privately.

Fifth, the guidelines in economic research recommend to use the most cost-effective alternative intervention currently available as a comparator (Drummond and Jefferson, 1996). Since in the detection of tubal pathology, the guidelines do not agree on a most cost-effective intervention, we decided to use the ‘do-nothing’ scenario as a comparator. By doing so, we could demonstrate the maximum gain to be obtained by testing for tubal pathology.

Sixth, we did not include intrauterine insemination (IUI) as treatment option but IVF instead. A recent randomised trial showed that IVF following three cycles of superovulation with Clomifene citrate, was more cost-effective than if IVF treatment was preceded by an additional three cycles of IUI and superovulation with Gonadotrophins, for unexplained subfertility (Reindollar et al., 2010). However a limitation of that study is that the mean and median duration of unexplained subfertility was not provided and the prognostic profile of the included couples cannot be extracted from the paper. This is important to interpret the effectiveness of IUI with superovulation versus expectant management (Steures et al., 2006, Custers et al. 2012). Because robust evidence of the effectiveness of IUI with superovulation in unexplained subfertility and without causing multiple pregnancies is still lacking, we omitted this treatment option in our scenarios (Battachyari et al. 2010 BMJ, Veltman-Verhulst et al., 2011).
Our study showed that routine diagnostic tubal patency tests in the fertility work-up are not cost-effective. However, apart from the information that tests supply for therapeutic decisions, tubal tests may have additional effects on patients’ health apart from the consequences of subsequent management decisions (Bossuyt and McCaffery, 2009; Lenhard et al., 2005). These additional effects, such as knowing the cause of the subfertility, being reassured that tubes are patent or anxiety provoked when the tests reveal bad news, were not considered in the present analysis. Obviously, adding the value of such information to the present scenarios could reduce the gap between the no-testing strategies and the strategies of tubal testing. In cost-effectiveness studies, the payer’s willingness to pay will influence the outcome of concern which in this study are live birth rates. Thus far this has not been addressed and should be the subject of further research. Be that as it may, when there is a need for information on the tubal status apart from management decisions, our analysis showed that a strategy based on HSG followed by laparoscopy if HSG shows bilateral occlusion, is superior over laparoscopy alone.

In view of all our findings we suggest the following for tubal patency tests in the fertility work-up, in women until 38 years and especially below the age of 30 years, expectant management and no diagnostic test for at least 12 months is justified and will reduce the number of unnecessary invasive diagnostic tests, complications and costs. An HSG followed by laparoscopy, if HSG shows bilateral occlusion, should be considered, if conception does not occur after expectant management and if a couple prefers fertility treatment other than IVF. In women with bilateral distal occlusion, HSG can be helpful to decide whether laparoscopic salpingostomy is preferable above or before IVF, although randomised evidence for this is lacking. In women from the age of 39 years and older, immediate treatment is the most cost-effective scenario. It is not to be expected that every couple is prepared to start immediately with IVF treatment or have the financial means to do so. The second best strategy is then to prove tubal patency by HSG and if the tubes are found to be open, couples can be counselled to choose between expectant management, intra uterine insemination or IVF, obviously taking the prognosis for natural conception into account (Bossuyt and McCaffery, 2009; Hunault et al., 2004; Steures et al., 2006). In some women sonographically visible bilateral hydrosalpinges are usually detected before tubal testing. In these women direct laparoscopy is advised and can be combined at the same time with salpingectomy, since it has been shown that this can increase IVF-outcome prior to IVF treatment (Johnson et al., 2010). In some women, severe endometriosis may have a negative impact on their tubal patency and natural pregnancy chances. This may be suspected after the initial fertility work-up or discovered at DL. Laparoscopic surgery and additional post-operative treatment may improve live birth rates in these women (Surrey et al., 2002), but often warrants a second laparoscopic procedure, because extra theatre time or even a referral to a specialised centre is required. Because of these logistic consequences and additional costs, our conclusions are unlikely to change.
In conclusion, the most cost-effective scenario is to perform no diagnostic tubal tests and delay IVF treatment for at least 12 months for women aged until 38 years and to perform no tubal tests and start immediate IVF treatment from the age of 39 years. If an invasive diagnostic test is planned, HSG followed by tailored treatment or a diagnostic laparoscopy in case HSG shows no tubal patency, is more cost-effective than diagnostic laparoscopy.
References


Veleva Z, Karinen P, Tomas C, Tapanainen JS, Martikainen H. Elective single embryo transfer with cryopreservation improves the outcome and diminishes the costs of IVF/ICSI. **Hum Reprod** 2009; **24**:1632-1639.

