Diagnostic tests for tubal pathology from a clinical and economic perspective

Verhoeve, H.R.

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.
Diagnostic tests in Reproductive Medicine

Harold Verhoeve
Fulco van der Veen
Ben Willem Mol

Reviews in Gynaecological and Perinatal Practice 2006, 6:20-25
Abstract

The aim of the fertility work-up is to exclude recognised causes of infertility and to distinguish those couples who have good spontaneous pregnancy prospects from those who have poor prospects. Information gathered by medical history, clinical findings and results of the diagnostic tests should help the clinician in counselling subfertile couples. The initial diagnostic tests for infertility should include a basal body temperature chart or midluteal phase progesterone level, a semen analysis and a test for tubal patency. Ovarian reserve tests can be used in selected cases. More research is needed in the predictive value of the outcome of diagnostic tests in reproductive medicine.
1. Introduction

Subfertility is defined as an inability to conceive after one year of regular unprotected sexual intercourse. One in six couples seek specialist care for this problem (Hull et al., 1985; Snick et al., 1997; De Kretser, 1997; Evers, 2002). In most clinics the fertility work-up consists of medical history taking and clinical examination, followed by diagnostic tests aimed at excluding recognised causes of infertility such as azoospermia, anovulation and bilateral tubal blockage. As a consequence, the initial diagnostic tests for infertility include evaluation of ovulation, a semen analysis and tests for tubal patency and function.

In this paper we will focus on existing data on diagnostic tests.

2. Assessment of ovulation

A history of a regular menstrual cycle indicates ovulation in more than 95% of women. Direct observation of ovulation is possible by monitoring follicular development sonographically until collapse is observed and fluid in the pouch of Douglas appears (Collins et al., 1991). This method however requires regular visits to the clinic and is costly. Ovulation can be predicted by measuring oestrogen or luteinizing hormone (LH) in the blood. Self-testing kits of LH in the urine are frequently used and easy to perform, but give a higher incidence of false negative results than blood tests (ESHRE Capri Workshop Group, 2000).

Confirmation if ovulation has occurred can be obtained by Basal Body Temperature (BBT), measurement of midluteal phase progesterone or endometrial dating through biopsy of the endometrium. A diagnosis should be made in cases of anovulation or oligo-ovulation. Thyroid stimulating hormone levels can point to thyroid disorder and Prolactine levels should exclude hyperprolactinaemia. These disorders can easily be corrected by treatment. Gonadotrophin levels in conjunction with oestrogen levels should be measured to assess if one is dealing with hypogonadotrophic hypogonadism, normogonadotrophic normogonadism or hypergonadotrophic hypogonadism. In case of normogonadotrophic normogonadism (WHO type II), androgen assays and ultrasound appearance of the ovaries are required for the diagnosis of polycystic ovary syndrome (PCOS) (Rotterdam PCOS consensus workshop group, 2004).

3. Assessment of Ovarian reserve

Natural fecundity and pregnancy rates decline with increasing age (Te Velde et al., 2002). This decrease in fecundability is related to a reduction in both quality and quantity of the primordial follicle pool (referred to as ovarian reserve) (Faddy et al., 1996). Several tests have been studied to screen for diminished ovarian reserve, mostly in relation to assisted reproductive technologies (ART).
3.1. FSH
Cycle day 3 serum (so called basal) FSH levels are widely used in many fertility centres. The FSH value is an indirect measure of serum inhibin-B and oestradiol that is being produced by a cohort of follicles and which concentrations result in feedback at the pituitary level. In an indexed cohort study (van Montfrans et al., 2000) the fecundity in sub-fertile patients below the age of 40 years, with elevated FSH levels (>10 IU/L) and regular ovulatory cycles, was not different from the fecundity in controls and both groups had comparable cumulative delivery rates. Another study showed that FSH did not have an independent predictive effect on pregnancy rates in a general subfertility population but should be interpreted in the light of other clinical variables like age (van Rooij et al., 2004). In a study of patients ≤ 41 years with elevated basal FSH levels it was shown that these patients had reasonable ongoing pregnancy rates, despite a considerable probability of cycle cancellation due to poor ovarian response, indicating that the quantity but not necessarily the quality of oocytes is diminished (van Rooij et al., 2003). A meta-analysis on the performance of basal FSH level in the prediction of poor ovarian response and failure to become pregnant after IVF showed that a possible clinical application of basal FSH refers to only a minority of patients with extremely high basal FSH levels (>20 IU/L) but should not be regarded as a useful routine test for the prediction of IVF outcome (Bancsi et al., 2003). In short, basal FSH is not a good screening test for ovarian reserve and pregnancy outcome in a general subfertility population and neither as a predictor for fertility outcome in patients undergoing ART. Inhibin B, like basal FSH, is not a useful screening test to predict diminished ovarian reserve (van Rooij et al., 2005).

3.2. AFC
Antral follicle count (AFC), defined as the number of follicles smaller than 10 mm in diameter in the early follicular phase, can be easily assessed by transvaginal ultrasound and was introduced to predict ovarian responsiveness in IVF treatment (Tomas et al., 1997). In a recent meta-analysis the performance of AFC in the prediction of poor ovarian response and pregnancy after IVF was assessed and compared to the performance of basal FSH (Hendriks et al., 2005). The AFC showed a significantly better performance than basal FSH in the prediction of poor ovarian response. In predicting non-pregnancy the predictive capacity of both tests was poor. In this study the performance of AFC was analysed independent of female age because of the lack of homogeneity between the available studies.
In a study of patients > 38 years multivariate logistic regression analysis showed the AFC to be a valuable test to assess the individual chance of pregnancy in IVF. The AFC was significantly associated with the occurrence of an ongoing pregnancy after IVF treatment, whereas age and basal FSH were not. However, pregnancy could not be excluded amongst those younger than 43 years of age (Klinkert et al., 2005).
3.3. AMH
Anti-Mullerian hormone (AMH) is produced by the granulosa cells of preantral and small antral follicles. In a longitudinal study of normal fertile women decline of serum AMH levels was associated with increasing age and as such this expresses a decline in reproductive function (van Rooij et al., 2005). In two studies decreased AMH levels were associated with poor ovarian response in IVF treatment but pregnancy outcome was not evaluated (Seifer et al., 2002; van Rooij et al., 2002). Like the AFC, AMH cannot identify poor IVF outcome with a certainty high enough to justify its use.

3.4. CCCT
Another ovarian reserve test that has been assessed is the clomiphene citrate challenge test (CCCT). This test evaluates the basal FSH level on cycle day 3 and the FSH response to clomiphene citrate administration from cycle day 5 to 9 (Navot et al., 1987). In comparison to basal FSH combined with AFC the predictive accuracy and clinical value in outcome of IVF of this test was not shown to be better and should not be used (Hendriks et al., 2005).

In general it can be concluded that AFC is a good predictor of ovarian response to gonadotrophins but not of pregnancy in young patients, whereas in the older patients, low AFC predicts poor response and poor pregnancy outcome in IVF treatment. It should therefore not be used as a routine test in the fertility work-up, but is useful in the older patient. Basal FSH and serum inhibin B levels should not be routinely used as a screening test for predicting ovarian response and pregnancy outcome. Anti-Mullerian factor, like the AFC could proof useful as a screening test in the older patient but this needs to be confirmed in further studies. The CCCT should not be used as a screening test for ovarian reserve.

4. Semen analysis
Azoospermia, repeatedly confirmed, implicates absolute infertility. When semen analysis adheres to the WHO reference values, i.e. a sperm concentration of 20 x 10^6/ml, 50% progressive motility (grades a and b) and 15% normal morphology according to strict criteria (WHO laboratory manual, 1999), the male is considered to be normal fertile. Despite the standardisation of semen analysis, the relationship between semen quality and biological fertility remains controversial (ESHRE Capri Workshop Group, 2004). The definition of normality is hampered by two problems. First, impaired male fertility can be compensated for by the female partner. Second, semen parameters are of predictive value in treatment independent pregnancy where the female has no reproductive dysfunction. This issue has been assessed in two case control studies that recommend values of 14.3-13.5 x 10^6/ml for concentration, 28%-32% for progressive motility and 5-9% for normal morphology (using strict criteria), respectively. Both studies thus showed comparable parameters in discriminating between fertile and subfertile
men. Morphology appeared to be the best discriminator between fertile and subfertile men (Ombelet et al., 1997; Guzik et al., 2001). A disadvantage of these case control studies is that they contrast subfertile men from fertile men, whereas the purpose of the fertility work-up is to distinguish those who have good spontaneous pregnancy prospects from those who have poor prospects. Bonde et al. performed such a study amongst first-pregnancy planners and found that use of WHO values might be inappropriate, as many men with subfertility would be inside the range of normality according to these values (Bonde et al., 1998).

In the presence of male subfertility the couple should be counselled which assisted reproductive treatment is most appropriate (IUI, IVF or ICSI), taking into account expected pregnancy rates, costs, patient discomfort and possible complications. The question is how to do this and which sperm parameters can select men for the best treatment option. Individual semen parameters like volume, concentration, and motility can be combined into the parameter of total motile sperm count (TMC), i.e., the total number of progressively motile spermatozoa present in the ejaculate. The TMC can be assessed directly from the ejaculate (prewash TMC) or after semen preparation (postwash TMC). The value of the postwash TMC in the fertility work-up could be in its prediction of patients that are unlikely to become pregnant after IUI and should therefore be advised IVF or ICSI. However, the TMC has only been assessed during the IUI cycles and not during the fertility work-up (van Weert et al., 2004).

5. Post-coital test

The post-coital test assesses the interaction between sperm and cervical mucus. A specimen of cervical mucus is obtained shortly before the expected ovulation and 8 to 16 hours after intercourse. The sample is examined under the microscope at 400x magnification and considered normal if at least 1 progressive motile spermatozoa is present in six high power fields (WHO laboratory manual, 1999). Controversies exist regarding technique, timing and interpretation of a normal test. In a systematic review of 11 observational studies the postcoital test showed a relative risk of pregnancy of 2, qualified as poor by the researchers, but potentially useful in clinical practice (Oei et al., 1995). The same group published a RCT which compared cumulative pregnancy rates between couples offered a postcoital test versus couples who were not offered this test as part of their fertility work-up (Oei et al., 1998). No significant differences were shown in their respective cumulative pregnancy rates (49%, 95% CI 42 to 55% in the PCT group versus 48%, 95% CI 42% to 55% in the control group). In this study no clear definition was given of a positive PCT result. The result did not alter the management of the patients and treatments were applied non-specifically and inconsistently. Anovulatory patients were included and treatment consisted of IUI with controlled ovarian hyperstimulation, not in a natural cycle. It is therefore not surprising that performing a PCT as part of the fertility workup had no effect on outcome.
If a well timed postcoital test is negative and no defined causes for infertility are found, cervical hostility is considered to be the cause of subfertility. Whether IUI is an effective treatment for cervical factor subfertility is not clear. Of five identified randomised trials the results were conflicting, but intrauterine insemination can give reasonable pregnancy rates in a natural cycle as well as in cycles with controlled ovarian hyperstimulation (Steures et al., 2004). The test can also be negative because of poor semen quality. In these cases, intrauterine insemination offers a benefit over timed intercourse both in natural and in controlled ovarian hyperstimulation cycles (Cohlen et al., 2000).

A retrospective study which examined the relation between test outcome and duration of subfertility showed that the postcoital test is an effective predictor of treatment-independent conception in the absence of defined female causes of infertility and provided that the duration of subfertility was short (Glazener et al., 2000). In couples with less than 3 years subfertility and a positive postcoital test, 68 % conceived within 2 years compared to 17% of the couples with a negative test result. In couples with a duration of subfertility of > 3 years the postcoital test result did not have any additional prognostic value. More recently, a prospective study (Hunault et al., 2005) assessed the validity of a prediction model for treatment-independent fertility outcome in couples (Hunault et al., 2004) and showed that the result of the PCT discriminated better between women who became pregnant and women who did not become pregnant than if the model was applied ignoring the result of the PCT.

However, although the PCT is not recommended as a routine diagnostic test (NICE, 2004), this recent evidence of the prognostic value of the PCT in the subfertile couple suggests that the PCT should be part of the fertility work-up. In order to reduce cost and burden for couples another recent study has shown that the outcome of the PCT can be predicted in about 50% of the patients without compromising the prediction of fertility prospects, by using information from the medical history and semen analysis (van der Steeg et al., 2004). Although in this study the PCT was performed by one gynaecologist, the level of experience in performing a PCT does not appear to have an effect on the predictive value of the PCT (Hunault et al., 2005).

6. Tubal assessment

Tubal obstruction is thought to be the cause of subfertility in 12 to 33% of subfertile couples (Collins et al., 1995). The prevalence depends on the reference population and has a tendency to increase from a primary-care to a tertiary-care level (Evers, 2002). Tubal patency tests usually are descriptive tests that visualise the cervical, uterine and fallopian morphology, and in case of laparoscopy or transvaginal hydrolaparoscopy, assess the pelvic cavity. They are not capable of assessing the tubal physiology and function that are necessary for successful conception (Croxatto et al., 2002). The medical history can be an indication of the presence of tuboperitoneal pathology (PID, septic abortion, endometriosis, pelvic surgery, chronic pelvic pain, dyspareunia) as can the clinical examination (pelvic tenderness, palpable mass).
6.1. Chlamydia Antibody Titre

*Chlamydia Trachomatis* is a major cause of pelvic inflammatory disease, leading to chronic abdominal pain, ectopic pregnancy and tubal factor infertility (Weström and Wolner-Hansen, 1993; Paavonen et al., 1999). The infection is asymptomatic in the majority of cases. The Chlamydia Antibody Titre (CAT) can be tested in the serum of women. CAT is based on the detection of a previous infection. Different serologic methods are available, but in a meta-analysis it was shown that enzyme immunoassay or (micro)–immunofluorescence (MIF/IF) performs best with a point estimate of 75% for both sensitivity and specificity (Mol et al., 1997).

In a recent study, in a tertiary centre, the tubal status of 1006 women were examined by laparoscopy and related to the serum CAT (Akande et al., 2003). This study showed a linear trend between serum CAT and the likelihood of tubal damage. Increasing antibody titres were quantitatively related to both presence of tubal damage and the severity of tubal damage. CAT performs equally well in the diagnosis of tubal pathology as a hysterosalpingogram (HSG) and can therefore be used as a triage before laparoscopy (Mol et al., 1997). If the CAT is negative (<1:64), tubal pathology at laparoscopy can be found in 7-12% of the patients (Land et al., 2003), severe damage is more likely if higher titres are found (Akande et al., 2003).

6.2. Hysterosalpingogram

HSG is a widely used test in the assessment of tubal patency. Interpretation of the results can be hampered by technical difficulties or tubal spasm.

The diagnostic performance of HSG was compared to laparoscopy and dye for tubal occlusion and peritubal adhesions in a meta-analysis (Swart et al., 1995). This meta-analysis showed a sensitivity of 65% (95% CI 50% to 78%) and specificity of 83% (95% CI 77% to 88%) for the diagnosis of tubal occlusion. The HSG is reasonably accurate for detecting proximal tubal occlusion, but not for distal tubal disease or peritubal adhesions (Swart et al., 1995; Mol et al. 1996). An abnormal result should be confirmed by laparoscopy, but with a normal result laparoscopy is likely to show tubal patency in 95% of cases (Mol et al., 1996). More importantly, one – sided occlusion detected using HSG was found to decrease the fecundity rate ratio slightly (FRR 0.80), whereas bilateral occlusion reduced fertility prospects more importantly (FRR 0.49) (Mol et al., 1999).

Centres performing HSG tend to use water-soluble contrast media (WSCM) rather than oil-soluble contrast media (OSCM) because of the potential side-effects of oil-soluble contrast media.

A recently updated Cochrane review (Johnson et al., 2005) found evidence of effectiveness of tubal flushing with OSCM if used therapeutically. OSCM increased the odds of pregnancy (OR 3.5, 95% CI 1.8 to 6.8) and live birth rate (OR 3.0, 95% CI 1.4 to 6.4) versus no intervention.
Limited evidence exists of an increase in the odds of live birth from tubal flushing with OSCM versus WSCM (OR 1.49, 95% CI 1.05 to 2.1) and must be interpreted cautiously, since this was based on two trials of which the higher quality trial showed no significant difference, justifying a further randomised trial comparing oil-soluble versus water-soluble media, where live birth is considered as the primary outcome (Johnson et al., 2005).

6.3. Laparoscopy and dye test
Laparoscopy and dye test allows the clinician to assess the pelvic cavity for presence of adhesions and endometriosis directly. Usually, this procedure is performed at the end of the fertility work-up, when a HSG or CAT has been assessed as being abnormal or sooner if the patient has comorbidity such as a medical history of PID, ectopic pregnancy or previous pelvic surgery. Laparoscopy can not be considered the gold standard for assessing tubal patency as presumed tubal obstruction may be due to differences in resistance between the tubes, spasm, or technical failure. This was illustrated in a cohort study comparing results of HSG and laparoscopy which implicated that 35% of the tubes that were found to be occluded at laparoscopy showed patency at HSG (Mol et al., 1999). In this study laparoscopy showed a better predictor for infertility than a HSG with a FRR of 0.51 for one-sided occlusion and FRR of 0.15 for two-sided occlusion. Findings of phimosis and/or adhesions of the tubes showed a FRR of 0.60, endometriosis grade I/II of 0.52, whereas no spontaneous pregnancies occurred in patients with endometriosis grade III/IV (Mol et al., 1999). Compared to HSG, laparoscopy has the advantage to assess if microsurgery is feasible in the presence of tubal pathology. If mild to moderate endometriosis is found, direct treatment appears to improve fertility prospects by 13% (Marcoux et al., 1997). This improvement could however be the result of adhesiolysis that was performed in the same setting, since the results of another randomised trial did not show a benefit from treating mild to moderate endometriosis on fertility outcome (Parazzini et al., 1999).

6.4. Selective salpingography and tubal catheterisation
If proximal tubal occlusion is suspected at HSG or laparoscopy, selective salpingography and tubal catheterisation can be performed under fluoroscopic guidance to assess whether proximal tubal blockage is caused by cornual spasm. In this procedure the tubes are cannulated and are flushed with contrast media. Tubal perfusion pressure, which can be measured using this test, appears to have prognostic value in predicting pregnancy (Papaioannou et al., 2004), but no randomised controlled trials have been published that assess the prognostic value.

6.5. Transvaginal hydrolaparoscopy
Transvaginal hydrolaparoscopy (THL) is a technique that makes use of micro-endoscopic instruments. A scope is entered through the posterior fornix of the vagina into the pelvic cavity. By means of aquaflotation, using instillation of warm saline, the tubo-ovarian structures are assessed. The procedure is usually combined with hysteroscopy and chromopertubation and
can be combined with fallopian tube palpation. The procedure can be performed using local anaesthesia or sedation and could be done in an outpatient setting. Findings at THL appear to be in agreement with laparoscopic findings (Watlerot et al. 2003). However, no studies have been published that assess the performance of THL as a predictor of treatment-independent fertility.

6.6. Hystero-contrast-sonography
Transvaginal hysterosalpinography contrast sonography (HyCoSy) combines the instillation of contrast agents into the uterine cavity with a transvaginal ultrasound and through this allows assessment of the uterine cavity, tubal patency as well as ovarian morphology. The procedure can be performed in an outpatient clinic and requires similar time to perform as the HSG. The performance of HyCoSy as a screening test for tubal infertility compares well with HSG with a high concordance rate. A diagnosis of occlusion requires additional laparoscopy for confirmation (Helpman et al., 2003, Dijkman et al. 2000).

In summary, to assess tubal patency, it is reasonable to start with screening for chlamydia antibodies. If negative, visual tests of tubal patency can be delayed for several months. If pregnancy has not occurred in these months, a visual test for tubal patency can be performed such as HSG or hystero-contrast-sonography. If these show tubal abnormalities, a laparoscopy and dye test is indicated. In patients with co-morbidity for tubal pathology or high chlamydia antibody titres, it is reasonable to proceed to laparoscopy as the initial investigation (ESHRE Capri Workshop Group, 2004, NICE). However, a rational approach for the assessment of tubal pathology can at present not be based on indisputable evidence.

7. Summary

Presence of ovulatory cycles can be easily assessed. Basal FSH level is not a good diagnostic test for ovarian reserve and pregnancy outcome in a general subfertility population and neither as a predictor of fertility outcome in patients with regular menstrual cycles undergoing ART. In the older patient, a low antral follicle count is a good predictor of poor ovarian response and poor pregnancy outcome in IVF treatment. More research is required into the predictive value of total motile count and other semen parameters in the presence of suboptimal semen quality. The post-coital test does have predictive value in pregnancy outcome and should be part of the fertility work-up, especially in couples with a duration of subfertility shorter than 3 years and normal semen parameters. More research into a rational approach for the assessment of tubal patency is needed.
References


CHAPTER 2


van Rooij IA, de Jong E, Broekmans FJ, Looman CW, Habbema JD, te Velde ER. High follicle-stimulating hormone levels should not necessarily lead to the exclusion of subfertile patients from treatment. *Fertil Steril* 2004; 81(6):1478-85.


