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.
Summary

Subfertility is customarily defined as failure to conceive after regular unprotected sexual intercourse for one year. The prevalence of subfertility is around 14%, affecting one in seven couples. It has been estimated that in around 10-30% of these couples, subfertility is due to tubal pathology, which includes tubal occlusion and pelvic adhesions due to previous infection, endometriosis or previous surgery. To assess the tubal condition, the fertility work-up in subfertile couples is usually concluded with tubal patency testing. Currently, laparoscopy is considered to be the best available test for diagnosing tubal abnormalities, but it has several drawbacks. First, it is an invasive surgical procedure that requires general anaesthesia, both of which carry associated risks. Secondly, as an expensive investigation that requires operating time and dedicated personnel, its availability is not unlimited. Therefore, triage is needed to limit the number of unnecessary laparoscopies, while maintaining a high diagnostic yield. Nationally and internationally a large practice variation exists in the timing of testing, the tests used, and the basis on which women are selected for tubal patency testing. While some guidelines advocate medical history taking to triage women for tubal testing, others advocate Chlamydia antibody testing, or testing in all women. Up to date, diagnostic test research on which these guidelines are based on, had been conducted in isolation from its clinical context and conventional meta-analyses, which are assumed to supply a maximum amount of data and provide the highest level of evidence are hampered by this examination of test accuracy isolated from medical history and clinical examination.

In this thesis we first introduced the approach of individual patient data meta-analysis in diagnostic and prognostic research in reproductive medicine. We started with the methodological issues of conventional meta-analyses and meta-analyses based on individual patient data. Hereafter, we performed IPD meta-analyses to reassess the accuracy of the Chlamydia antibody test (CAT) by using its continuous test results and to investigate the influence of patient characteristics on the diagnostic accuracy of hysterosalpingography (HSG). Next, we estimated the best strategy to diagnose tubal pathology, using information from medical history, clinical examination and tubal patency test results. Finally, in a fourth IPD meta-analysis, we investigated the predictive capacity of patient characteristics, HSG and diagnostic laparoscopy to predict natural conception.

Chapter 1
This chapter gives an outline and description of the objectives of this thesis.

Chapter 2
This chapter presents an opinion article that describes the advantages and disadvantages of conventional systematic reviews and meta-analyses in reproductive medicine. It also addresses the opportunities of systematic reviews and meta-analyses using individual patient data, based on three clinical examples of current studies. Systematic reviews and accompanying
Chapter 9

Meta-analyses are assumed to be the cornerstones of evidence based medicine. They summarize clinical evidence and provide summary estimates of treatment effect or diagnostic test accuracy. However, as they can only summarize aggregated data, they suffer from a number of drawbacks. Besides the possibility of publication bias and poor reporting in the original studies, more methodological problems exists, such as different end points of follow-up in the included studies than those of primary interest or the impossibility to distinguish between patients with different clinical profiles. Therefore, the clinical value of conventional meta-analyses is limited. Systematic reviews and meta-analyses using individual patient data offer a promising, but challenging approach, by which treatment effects and diagnostic accuracy can be estimated at the level of relevant patient subgroups. Also, direct contact and collaboration with the original authors improves data collection, prevents publication bias and combats poor reporting. Three clinical examples were described to illustrate the possible advantages of IPD meta-analysis. The first example is an IPD meta-analysis on the effectiveness of single versus double embryo transfer (SET resp. DET) in IVF treatment. In this meta-analysis summary odds ratios for live birth rate and multiple birth rate were estimated, adjusted for both patient characteristics as well as for laboratory characteristics that may influence the effectiveness of therapy. This type of data may lead to a valid identification of patient subgroups that will benefit from SET or will be harmed by DET. The second example describes an IPD meta-analysis on the diagnosis of tubal pathology that is described in detail in chapters 4 to 7 of this thesis. Using this individual patient data allowed for studying several sequences of tests, for integration of different definitions of tubal pathology and patient characteristics based on both clinical history and physical examination, as well as for the use of different CAT assays and continuous test results. In the third and last example, the aim of the IPD meta-analysis was to prove or disprove that addition of patient specific information, like female age, would improve the predictive value of ovarian reserve tests in predicting response to ovarian stimulation and pregnancy outcome. This might enable the creation of a predictive tool that identifies subgroups of women with specific pregnancy chances, based on their clinical profile.

Chapter 3

This chapter describes a protocol in which the general methodology of IPD meta-analysis is described and a framework for future IPD meta-analyses is provided. Four clinical problems in the field of obstetrics, gynaecology and reproductive medicine are used to illustrate this new analysis approach. 1. Diagnosis of endometrial cancer in women with postmenopausal bleeding (PMB). 2. Prediction of preterm birth. 3. Diagnosis of tubal pathology in subfertile women. 4. Assessment of ovarian response in women undergoing in vitro fertilisation (IVF). Previously, systematic reviews of each of the four clinical topics had been performed and relevant primary research was identified. For the performance of an IPD meta-analysis, the first step is to approach the first authors of the included studies and ask them to share their original, complete datasets, provided with a clear legend. These acquired datasets are
assessed for validity and completeness and the study quality is reported according to the STARD statement. If the variables are compatible the original data is merged and a study identification variable is added to reflect the stratified nature of the pooled dataset. Based on this dataset, series of analyses can be performed, including a systematic comparison of the results of the IPD meta-analysis with those of a conventional meta-analysis and the development of multivariable models both for clinical history and examination alone, as well as for various combinations and sequences of relevant patient characteristics with additional test results. The clinical “end products” of these IPD meta-analyses will be prediction rules for each of the four clinical problems: women with PMB, women at risk for preterm birth, women suspected of having tubal pathology, and women starting with IVF. The results can be made available through simple scoring charts as well as logistic regression models. The latter will become accessible through web applications at which doctors can enter relevant data of the individual patient, to offer the best diagnostic strategy and management for the individual patient. The described clinical topic on the diagnosis of tubal pathology in subfertile women is elaborated in this thesis.

Chapter 4
The Chlamydia antibody test (CAT) shows considerable variations in reported estimates of test accuracy in diagnosing tubal pathology, partly because of the use of different assays and cut-off values. The aim of this study was to reassess the accuracy of CAT for three different assays, using the continuous test results of individual patients. After a systematic literature search, authors of primary studies that used micro immunofluorescence tests (MIF), immunoassay tests (IF) or enzyme-linked immunosorbent assay tests (ELISA) in the diagnosis of tubal pathology, were approached to share their original data. From 14 primary studies, data of 6,191 women were obtained, containing continuous CAT results of 3,453 women that were suitable for the analysis. After multiple imputation was performed to correct for missing laparoscopies and reducing verification bias, the overall accuracy of the three different essays was estimated. ROC curves were generated by ROC-analysis based on the predicted probabilities from a random effects logistic regression model to adjust for the heterogeneity across studies. For ELISA, IF and MIF the areas under the ROC curves were 0.64, 0.65 and 0.75, respectively (P value < 0.001) for any tubal pathology and 0.66, 0.66 and 0.77, respectively (P value = 0.01) for bilateral tubal pathology. The accuracy of MIF showed a moderate ability to discriminate between women with and without tubal pathology, whereas ELISA and IF only showed poor discriminative ability, indicating that the micro immunofluorescence test (MIF) is superior in the assessment of tubal pathology and should therefore be the test of first choice in the initial screen for tubal pathology.

Chapter 5
The sensitivity and specificity of hysterosalpingography (HSG) has been estimated by conventional meta-analysis to be 65% and 83%. However, the impact of patient
characteristics on this accuracy of HSG is still unknown. The aim of this study was to assess whether the accuracy of HSG is associated with patient characteristics. Authors of primary studies reporting on the accuracy of HSG using findings at laparoscopy as the reference were approached and the obtained individual patient data were used in an IPD meta-analysis. Data of seven primary studies were obtained, containing data of 4,521 women. The pooled sensitivity and specificity of HSG were 53% and 87% for any tubal pathology and 46% and 95% for bilateral tubal pathology. The possible association between patient characteristics such as female age, duration of subfertility and a clinical history without risk factors for tubal pathology and the accuracy of HSG was assessed using a random intercept logistic regression model. This model showed that the sensitivity of HSG was not associated with patient characteristics, except for women without risk factors for tubal pathology, consisting of no previous PID and a negative CAT result. In these women the sensitivity of HSG was 38% for any tubal pathology, compared to 61% in women with risk factors ($P = 0.005$). For bilateral tubal pathology, these rates were 13% versus 47% ($P = 0.01$), which might probably due to laparoscopic artefacts. The specificity of HSG was relatively high and very stable across all subgroups. It could be concluded that HSG performs equally in women with different profiles.

Chapter 6

Tubal patency tests are routinely performed in the diagnostic work-up of subfertile patients, but it is unknown whether these diagnostic tests add value beyond the information obtained by medical history taking and findings at physical examination. In this study an individual patient data meta-analysis was performed to assess this question. After a systematic literature search, authors of primary studies for datasets containing information on patient characteristics and results from tubal patency tests, such as Chlamydia Antibody Test (CAT), hysterosalpingography (HSG) and laparoscopy were approached to share their data. From four studies reporting on 4,883 women, data could be obtained. Logistic regression modelling was performed to create models that predict tubal pathology from medical history and physical examination alone, as well as models in which the results of tubal patency tests are integrated in the patient characteristics model. From the available patient characteristics, duration of subfertility, previous pregnancies, previous pelvic inflammatory disease (PID), pelvic surgery and a history of Chlamydia infection were qualified for the patient characteristics model. This model showed an AUC of 0.63 (95% CI 0.61 to 0.65). For any tubal pathology, addition of HSG significantly improved the predictive performance to an AUC of 0.74 (95% CI 0.73 to 0.76) ($p < 0.001$). For bilateral tubal pathology, addition of both CAT and HSG to patient characteristics increased the predictive performance to an AUC of 0.76 (95% CI 0.74 to 0.79) and is therefore the best diagnostic strategy to diagnose bilateral tubal pathology.
Chapter 7
To provide optimal management for subfertile couples, it is essential to calculate the chances of natural conception. Previous studies on the predictive capacity of HSG and laparoscopy estimated this capacity without taking into account patient characteristics, thereby ignoring possible relations between patient profile and tubal status. The aim of this study was to assess the predictive capacity of tubal patency tests for natural conception in women with different profiles and to investigate the added value of these tests over patient characteristics obtained from clinical history and physical examination. After a systematic literature search, data of 6,326 women could be obtained for the analyses. For women with unilateral tubal pathology at HSG or laparoscopy, increasing age lead 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 all women. Multivariable models were built to assess the prognostic value of patient characteristics and the added value of HSG and laparoscopy. A model based on female age, duration and type of subfertility, BMI, 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, 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. After taking into account patient characteristics, findings at HSG or laparoscopy still had significant prognostic importance. However, laparoscopy had no added value when HSG was already performed. It was concluded that natural conception can be predicted from patient characteristics and hysterosalpingography results and was most reduced in women with bilateral tubal pathology at HSG.

Chapter 8
In this chapter the findings of this thesis are discussed, clinical implications are given and future research recommendations are made.

Chapter 10
This chapter provides an epilogue in which the results of three different theses on tubal pathology from our study group are integrated.