Universiteit
Leiden
The Netherlands

Evaluation of diagnostic and therapeutic strategies in reproductive medicine: studies on hysterosalpingography and assisted conception

Perquin, D.A.M.

Citation
Perquin, D. A. M. (2007, October 18). *Evaluation of diagnostic and therapeutic strategies in reproductive medicine: studies on hysterosalpingography and assisted conception*. Retrieved from https://hdl.handle.net/1887/12385

Version:	Corrected Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/12385

Note: To cite this publication please use the final published version (if applicable).

Denise Perquin

Evaluation of diagnostic and therapeutic strategies in reproductive medicine: studies on hysterosalpingography and assisted conception

Evaluation of diagnostic and therapeutic strategies in reproductive medicine: studies on hysterosalpingography and assisted conception

ISBN 978-90-9022338-4

Cover: Jan Snoeck
Printed by: Pasmans Offsetdrukkerij BV, Den Haag

Financial support for the printing of this thesis was provided by Ferring BV, MerckSerono, Organon Nederland BV, Abbott BV, Bayer BV, LIOBA Stichting, J.E. Jurriaanse Stichting and Wetenschapsfonds Medisch Centrum Haaglanden.

© 2007 D.A.M. Perquin

Evaluation of diagnostic and therapeutic strategies in reproductive medicine: studies on hysterosalpingography and assisted conception

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus prof.mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 18 oktober 2007
klokke 13.45 uur

door

Denise Arlette Maria Perquin

geboren te Dordrecht
in 1968

Promotiecommissie

Promotor:	Prof. Dr. F.M. Helmerhorst
Co-promotores:	Dr. A.J.M. de Craen
	Dr. P.J. Dörr (Medisch Centrum Haaglanden, Den Haag)
Referent:	Prof. Dr. M. Dhont (Universitair Ziekenhuis Gent, België)
Overige leden:	Prof. Dr. H.H.H. Kanhai
	Prof. Dr. J.A. Land (Universitair Medisch Centrum Groningen)
	Prof. Dr. J.P. Vandenbroucke
	Prof. Dr. F.J. Walther

Aan mijn ouders

Voor Maarten, Berend en Wouter

Contents

Chapter 1	General introduction	9
Chapter 2	Routine use of hysterosalpingography prior to laparoscopy in the fertility workup: a multicenter randomised controlled trial. *Hum Reprod 2006;21:1227-31*	29
Chapter 3	Experiences of subfertile women undergoing hysterosalpingography or laparoscopy and dye. *Submitted for publication*	41
Chapter 4	The value of Chlamydia trachomatis-specific IgG antibody testing and hysterosalpingography for predicting tubal pathology and occurrence of pregnancy. *Fertil Steril 2007;88:224-6*	53
Chapter 5	Difficulties in recruitment for a randomized controlled trial involving hysterosalpingography. *Reprod Health 2006;3:5*	61
Chapter 6	Perinatal outcome of singletons and twins after assisted conception: a systematic review of controlled studies. *BMJ 2004;328:261*	69
Chapter 7	General discussion	101
Chapter 8	Summary & Samenvatting	115
Addendum		123
Dankwoord		131
Curriculum Vitae		135

Chapter

1

General introduction

1. General introduction

Two studies on evaluation of diagnostic and therapeutic strategies in reproductive medicine are presented in this thesis. The first subject is about the diagnostic strategy and describes the results of a pragmatic multicenter randomized controlled trial comparing fertility work-ups with or without hysterosalpingography (HSG) to assess the value of HSG prior to laparoscopy and dye in a routine clinical setting. The second subject focusses on therapy in reproductive medicine and is based on a study on assisted conception to clarify whether adverse perinatal outcome is related to assisted conception or other maternal risk factors and whether this is for both singleton and twin pregnancies.
First all backgrounds are given in this chapter on the history and evolution of the Cochrane Collaboration followed by an overview of diagnostic and therapeutic strategies in reproductive medicine. At the end of the chapter an outline of this thesis is summarized.

1.1 Evidence based medicine in Obstetrics, Gynaecology and Reproductive Medicine

In 1979 the epidemiologist Archie Cochrane (1908-1988) published an essay, in which he suggested that "*It is surely a great criticism of our profession that we have not organised a critical summary, by specialty or subspecialty, adapted periodically, of all relevant randomised controlled trials"*[1]. He also designated Obstetrics the least scientific medical specialty although his words were also applicable for Gynaecology and the emerging field of Reproductive Medicine (www.cochrane.org/docs/cchronol.htm). This statement, combined with direct funding from the World Health Organisation and the English Department of Health, led to a register of controlled trials in perinatal medicine, which was located in the National Perinatal Epidemiology Unit, Oxford. Based on this register an international collaboration established an institute that enabled physicians and scientists to prepare systematic reviews of controlled trials on intervention in pregnancy, childbirth, and the neonatal period. In 1988 and 1989 the collaborative work was brought together in publications called *'The Oxford Database of Perinatal Trials'* (ODPT)[2]*, 'Effective care in Pregnancy and Childbirth (ECPC)*[3] *and 'a guide to effective care in Pregnancy and Childbirth' (GECPC)*[4] composed by Ian Chalmers, Murray Enkin, and Marc Keirse. This monumental work was the launch of the Cochrane Collaboration established by Ian Chalmers in 1993. The Cochrane Collaboration is an international non-profit and independent organisation, dedicated to making up-to-date, accurate, and readily available summaries on the effects of healthcare interventions. It produces and disseminates systematic reviews of healthcare interventions and promotes the

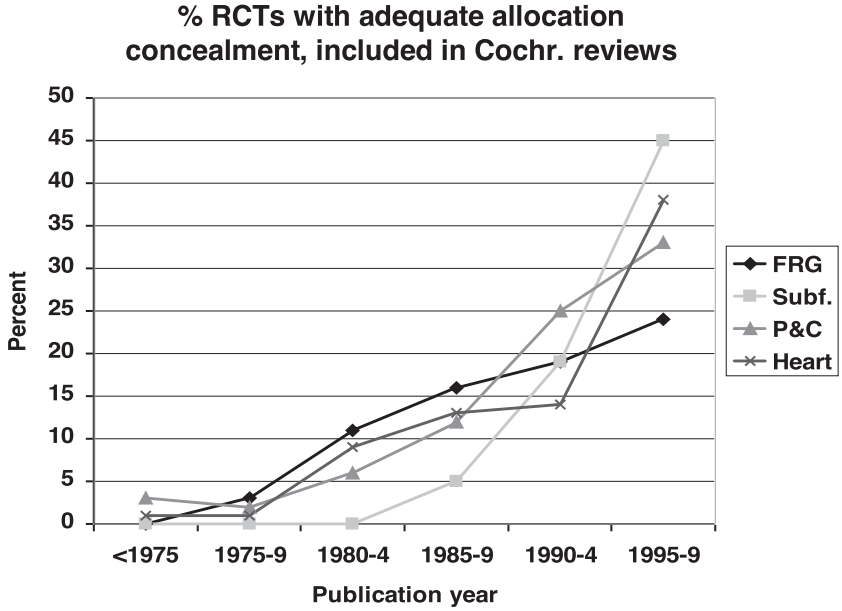


Figure 1. Proportion of randomized controlled trials with adequate allocation concealment that are included in published Cochrane reviews in four content areas, by years. FRG, fertility regulation; Subf, subfertility; P&C, pregnancy and childbirth, and heart (Helmerhorst *et al.,* Contraception 2006)

search for evidence by clinical trials and other studies of interventions. As part of the Cochrane Collaboration, the specialty of Obstetrics & Gynaecology registered their subspecialties: the Pregnancy & Childbirth group and the Subfertility group (1992), the Neonatal group (1993), the Menstrual Disorders group (1995), the Gynaecological Cancer group (1996) and the Fertility Regulation group (1997). Together with the development of systematic reviews in this period of time, not only the quantity, but also the quality of randomised controlled trials, the basis of the Cochrane systematic reviews, increased greatly[5].

Allocation concealment is a technique used to prevent selection bias by concealing the allocation sequence from those assigning participants to intervention groups, until the moment of assignment[6]. Allocation concealment prevents researchers from (unconsciously or otherwise) influencing which participants are assigned to a given intervention group. In 1994, Schulz *et al*.[7] reported that only 23% of the randomised controlled trials published in obstetrics and gynaecology journals contained information about adequate allocation concealment.

Since then the proportion of included trials with documentation of adequate allocation concealment has been progressively increasing in the subspecialties subfertility, perinatology, and fertility regulation[5].

1.1.1 Randomised controlled trial

The randomised controlled trial is by large the best tool we have, to demonstrate the possible benefits of an intervention, especially when a small difference in outcome between the two arms is expected. In 1999, Haynes[8] referred to Archie Cochrane for the differentiation between the three concepts related to testing healthcare interventions: efficacy, effectiveness and efficiency. "Efficacy is the extent to which an intervention does more good than harm under ideal circumstances ("Can it work?"). Effectiveness assesses whether an intervention does more good than harm when provided under usual circumstances of healthcare practice ("Does it work in practice?"). Efficiency measures the effect of an intervention in relation to the resources it consumes ("Is it worth it?"). Trials of efficacy and effectiveness have also been described as explanatory and management trials, respectively, and efficiency trials are more often called cost effectiveness or cost benefit studies."

1.1.2 Observational studies

In his commentary on 'Balancing benefits and harms in health care', Vandenbroucke[9] stood up "for the world of pharmacoepidemiology, which mainly investigates the harms of the same treatments" by showing that "observational data on harm should complement systematic reviews of benefit". Observational data are considered as intrinsically problematic. Evidence hierarchies have been proposed with the randomised trial at the apex and the observational studies dangling far behind it[10]. However, for the identification of unintended, infrequently occurring adverse effects, only observational methodology, such as case control and cohort studies, can help us. Reconfirmation of these results surfaced by observational studies can be established by randomised controlled trials, as we have seen for venous thromboembolism in the Women's Health Initiative studies on hormonal replacement therapy[11].

1.1.3 Background of thesis

There are many arguments why HSG should or should not be part of the standard diagnostic fertility work-up based on its diagnostic value, conception rates afterwards, pain, and costs. Therefore, a pragmatic multicenter randomised controlled trial comparing fertility work-ups with or without HSG was conducted at Leiden University Medical Center in collaboration with Westeinde Hospital (The Hague) and Groene Hart Hospital (Gouda) to assess the value of HSG prior to laparoscopy and dye in a routine clinical setting. A randomised comparison of two infertility protocols that

were identical except for inclusion and exclusion of HSG was conducted. Thus, the entity of a diagnostic method together with the decision for a therapy, was tested with the primary, hard clinical outcome 'pregnancy'. This procedure reflects the situation in the clinic and is called a pragmatic trial. It was this approach that was considered applicable for testing the routine use of hysterosalpingography prior to laparoscopy in the fertility workup, being the first subject of investigation as presented in this thesis.

In 1978, Louise Brown was born as first IVF baby[12]. Since then assisted reproductive technologies have improved with increasing pregnancy rates. However, outcomes other than live birth rates have become more important. As early as 1985, the Australian In Vitro Fertilisation Collaborative Group has shown increased adverse perinatal outcomes (such as preterm birth and low birthweight) in singleton pregnancies conceived after assisted conception compared to naturally conceived pregnancies[13]. But it was unclear by what extent the adverse perinatal outcomes were related to assisted conception or to confounding factors, such as maternal age and parity. After 1985, several cohort studies have confirmed the findings from Australia[14-16] but some studies found an opposite trend[17,18]. Moreover, for twin pregnancies the general consensus, with few exceptions[19-21], seemed to be that twin pregnancies conceived after assisted conception have outcomes that are either similar to or slightly better than those conceived naturally[14,17,22-25].

However, no systematic reviews were available to clarify whether perinatal outcome was related to assisted conception or other maternal risk factors, and whether this was for both singleton and twin pregnancies. We performed a systematic review of 25 controlled studies published between 1985 and 2002, to examine whether there are differences in outcome between assisted and natural conceptions and whether they apply to both singleton and twin pregnancies. This subject is the second investigation of this thesis.

1.2 Diagnostic strategies in reproductive medicine

1.2.1 Diagnostic fertility work-up

A couple is proclaimed subfertile when they have one year of regular, unprotected intercourse without conception. The prevalence of subfertility in the Netherlands is about 10-17% depending on the age of women[26]. The basic diagnostic investigations for both men and female in the fertility work-up as recommended in the National Institute for Clinical Excellence (NICE) guidelines[27], exist of history taking, clinical examination, semen analysis, ovulation detection and assessment of tubal damage. Here we discuss several techniques for assessing tubal function, laparoscopy and dye

as the reference standard, hysterosalpingography (HSG) and Chlamydia antibody testing (CAT).

1.2.2 Assessment of tubal pathology

Laparoscopy and dye

The first gynaecologic laparoscopy was performed by Raoul Palmer in 1944. About 25 years later, in the late 1960s, the diagnostic laparoscopy became a routine procedure in gynaecological practices. Laparoscopy and dye, also called dye hydrotubation, dye insufflation, dye pertubation, chromopertubation, or chromolaparoscopy, allows direct visualisation of the fallopian tubes. The laparoscopy enables assessment of tubal patency, peritubal adhesions, and endometriosis. Laparoscopy is applied under general anaesthesia and it is an invasive procedure with risks between 0.06% and 0.20% for vascular, intestinal, and urological injuries[28]. Anaesthetic complications and methylene blue toxicity have also been reported, but are extremely rare[29].

Laparoscopy is generally regarded as the definitive test for the evaluation of tubal pathology. Additionally, laparoscopic treatments such as coagulation of endometriosis grade I/II, adhesiolysis, or cystectomy, can be performed in the same session. There is a possible fecundity enhancing effect after laparoscopic coagulation of minimal and mild endometriosis plus adhesiolysis (www.eshre.com). Although laparoscopy and dye is the reference standard in the diagnosis of tubal pathology according to World Health Organisation (WHO) criteria[30], it is not beyond criticism[31]. Moreover, laparoscopy sometimes returns false positive results, meaning that women sometimes do conceive after bilateral obstruction according to laparoscopy.

The risk of having peritubal abnormalities or endometriosis in subfertile women who are asymptomatic, have normal physical examination, and have bilateral tubal patency at HSG, is low. In these women laparoscopy could be delayed or abandoned because only in a small minority of women severe tubal pathology will be revealed[27,32]. Moreover, women with co-morbidities, such as endometriosis or history of pelvic inflammatory disease or ectopic pregnancy should be offered a laparoscopy and dye early in the fertility work-up, according to RCOG and NICE guidelines[27,32]. Whether flushing of the fallopian tubes enhances fertility by laparoscopy and dye is not known. One randomized controlled trial showed no difference in pregnancy rate between laparoscopic chromopertubation with oil contrast medium or water soluble contrast medium[33].

Hysterosalpingography

In 1910, Rindfleisch described the visualisation of the uterine cavity by delivering

a bismuth solution through the vagina[34]. In 1920, Rubin introduced the tubal insufflation test using carbon dioxide to investigate tubal patency[35]. In 1925, Heuser performed the first hysterosalpingography with oil soluble contrast media[36].

At hysterosalpingography (HSG) contrast dye is injected through the cervical canal into the uterine cavity. Subsequent X-ray imaging visualizes the uterine cavity and the fallopian tubes. HSG is being used predominantly in the fertility work-up for the evaluation of uterine cavity and fallopian tubes. Other indications for HSG include the evaluation of uterine anomalies in women with a history of recurrent pregnancy loss and the postoperative evaluation of tubal patency or blockage.

Because patients may experience cramping during the HSG, women are recommended to take a nonsteroidal anti-inflammatory drug (NSAID) before the procedure to reduce discomfort.
Contraindications to HSG are iodine allergy[37] and active pelvic infection. Prophylactic antibiotics should be considered for patients having history of pelvic infections, positive result Chlamydia antibody testing[38] or tubal pathology assumed at HSG. However, it is not clear from the literature whether antibiotic prophylaxis following transcervical intrauterine procedures protects against infection. The risk of infection after hysterosalpingography is about 1-3%[39].

Traditionally HSGs were performed with oil-soluble contrast media. These oil-soluble contrast media have gradually been replaced by water-soluble contrast media. There are a number of reasons for using water-soluble contrast media in stead of oil-soluble contrast media:

1. Water soluble contrast media are generally cheaper than oil-soluble contrast media.
2. Water soluble contrast media have better imaging of the tubal mucosal folds and ampullary ruggae than oil-soluble contrast media[40].
3. Water soluble contrast media have a low viscosity which results in quick filling of the fallopian tubes and dispersion in the pelvis, therefore a delayed film is performed within one hour. In contrast with oil contrast medium what is more viscous and disperses very slowly in the pelvis, therefore a delayed film should be performed after 24 hours.
4. Less severe adverse events have been reported with water-soluble contrast media, inclusing pulmonary embolus, cerebral embolus, granuloma formation, anaphylaxis, and even death as result of intravasation. Since the introduction of low-viscosity oil contrast media, these serious pulmonary complications due to oil-soluble contrast media are extremely rare nowadays[41]. However, a significant

lower probability of immediate pain, prolonged pain, and post-procedure pain occurs with oil-soluble contrast media compared to water-soluble contrast media[42].

Most fertility centers have moved towards the use of HSG with water soluble contrast media rather than oil-soluble contrast media for the beneficial reasons outlined above[43].

Whether oil-soluble or water-soluble contrast media might have a fertility enhancing effect, has been under debate for half a century. A meta-analysis of randomised controlled trials showed a significant increase in pregnancy and live birth rates by flushing the tubes with oil-soluble media when compared with no intervention[42]. No such data exist for water-soluble contrast media versus no intervention. There are no randomized trials of good quality for the comparison of tubal flushing with oil-soluble media versus tubal flushing with water-soluble media, regarding the odds of pregnancy or live birth rates. The underlying mechanism of the possible fertility enhancing effect of oil-soluble media, whether this is a 'tubal flushing' phenomenon, an effect on the intraperitoneal environment, or an implantation enhancing effect on the endometrium, is still unclear[44,45].

HSG plays a role in the diagnostic evaluation of abnormalities relating to the uterus and fallopian tubes:

1. Uterine abnormalities that can be detected with HSG include congenital malformations (like uterus unicornuate, septate, bicornuate, and didelphys), polyps, submucous leiomyomas and adhesions. Uterine cavity abnormalities with a prevalence of about 10-15% in subfertile women can be visualized by HSG. Other methods for assessing the uterine cavity are transvaginal sonography, saline infusion sonography and hysteroscopy. Hysteroscopy (as gold standard) is generally used for confirmation and treatment of uterine cavity abnormalities found on HSG. It is important to bear in mind that the effectiveness of surgical treatment of uterine abnormalities to enhance pregnancy rates has not yet been established[27].

The endovaginal ultrasound has been demonstrated as a reliable alternative for HSG in detecting uterine cavity pathology[46-47]. Ultrasonography compared to hysterosalpingography is less invasive and more easily performed in a short period without roentgen radiation. It can also be applied in the visualization of pelvic pathology such as endometriosis or ovarian pathology.

2. The most important role of HSG is the assessment of tubal pathology. Tubal

abnormalities that can be detected include proximal tubal occlusion, salpingitis isthmica nodosa, hydrosalpinx, and peritubal adhesions. The tubal abnormalities seen at HSG can be either a congenital anomaly or due to cornual spasm, occlusion, or infection. At radiology, tubal spasm can not be distinguished from proximal tubal occlusion.

Laparoscopy and HSG

Several authors have compared HSG with laparoscopy as the reference standard for tubal patency. A meta-analysis of 20 studies between 1968 and 1994 concluded that while a normal HSG without tubal blockage will not rule out tubal disease, an abnormal HSG with tubal blockage is good at ruling in tubal pathology[48]. Therefore, with point estimates of 65% of sensitivity and 83% of specificity, HSG has moderate sensitivity and specificity in predicting tubal patency compared to laparoscopy and dye[48]. This meta-analysis also showed a very poor performance of HSG for predicting peritubal adhesions.

In daily practise many clinicians perform a laparoscopy soon after HSG, if at HSG tubal pathology is assumed. If HSG showes no tubal pathology and the woman does not conceive within six months, a laparoscopy and dye followes after six month. This is according to the RCOG and NICE guidelines[27,32] in which HSG has been considered as a screening test for tubal patency in low-risk couples.

Further, considerable variability in the interpretation as well as in clinical management of HSG abnormalities, has been shown among practitioners[49,50]. As might be expected, the probability of treatment-independent pregnancy is best when HSG reveals absence of tubal occlusion, substantially lower when there is two-sided occlusion, and only slightly reduced in the case of one-sided occlusion at HSG[31,51,52].

HSG with water-soluble contrast media remains the only trustworthy method for examining tubal mucosal fold configuration, especially when tubal microsurgery (e.g. salpingostomy) is considered[53], albeit that HSG then can be better carried out after laparoscopy.

Chlamydia trachomatis antibody testing

Chlamydia trachomatis is the most common sexually transmitted disease worldwide and remains asymptomatic in up to 80% of women[54]. Thus, women may remain untreated for years leading to long-term sequelae such as ectopic pregnancy and tubal subfertility. Pelvic inflammatory disease caused by *Chlamydia trachomatis* is

the major cause of tubal subfertility (http://www.who.int/vaccine_research/diseases/soa_std/en/index.html) and results in distal tubal obstruction and pelvic adhesions. Since the association between *Chlamydia trachomatis* antibodies in serum and tubal pathology was noticed in 1979[55], Chlamydia trachomatis antibody testing (CAT) has been introduced for large-scale screening purposes in the fertility work-up. Although both laparoscopy and hysterosalpingography are frequently used in the diagnostic fertility work-up of tubal pelvic damage, they are costly and invasive procedures. CAT is nowadays commonly used to assess the risk of tubal pathology in subfertile women in an inexpensive and non-invasive way. Because of the high negative predictive value of CAT in subfertile women (85–90%)[56], the presence of tubal pathology in patients with a negative CAT is unlikely. The positive predictive value of CAT in subfertile women is lower than the NPV and ranges from 30 to 65%[56]. The results reported on the diagnostic accuracy of CAT are heterogenous because of differences in CAT tests, threshold levels for a positive test, reference standard and definition of tubal pathology used. However, the simplicity and limited inconvenience of CAT serology enables its wide use in infertility practise to screen for tubal damage[57].
CAT and HSG provide risk estimates of tubal pathology prior to laparoscopy, but the diagnosis tubal pathology can only be made with laparoscopy and dye. Until now, there is no consensus on which tests should be used routinely for the diagnostic work-up of tubal pathology.

New diagnostic procedures
There are other new diagnostic investigations for assessing tubal damage, like hysterosalpingo-contrast-sonography or transvaginal laparoscopy. It is beyond this thesis to explain these procedures in more details.

1.3 Therapeutic strategies in reproductive medicine

1.3.1 Types of assisted reproductive treatments (ART)

In the last two decennia, assisted reproductive treatments (ART) such as ovarian stimulation, intra-uterine insemination (IUI), in vitro fertilization (IVF), and intracytoplasmatic sperm injection (ICSI), have become widely applied methods in the treatment of subfertile couples. These treatments will be explained briefly.

Ovarian stimulation
Two methods of ovarian stimulation should be distinguished. Mono-follicle development is attempted in anovulatory women and secondly, ovarian hyperstimulation is usually used to treat ovulatory women with the aim to induce

development of more than one dominant follicle to facilitate assisted conception. Ovarian hyperstimulation is sometimes combined with intra-uterine insemination, and nearly always before in vitro fertilization and intracytoplasmatic sperm injection.

Intra-uterine insemination (IUI)
Intra-uterine insemination means artificial insemination of in vitro selected motile spermatozoa. IUI treatment is widely used in subfertile couples with unexplained subfertility problems, minimal to mild endometriosis and mild male factor fertility problems[27]. IUI is often offered before IVF because IUI is less invasive, requires less intensive monitoring, and is associated with lower risks of hyperstimulation and multiple pregnancy.

The IUI procedure implies monitoring follicular growth development (1-3 follicles) by ovarian ultrasound. IUI can be performed with or without controlled ovarian hyperstimulation to increase the number of follicles slightly. Insemination is performed around the time of ovulation. The processed semen is inseminated directly into the uterine cavity through a small catheter bypassing the cervical canal, after which fertilization takes place in vivo.

In vitro fertilization (IVF)
In vitro fertilization is an assisted reproductive technique in which fertilization of the oocyte takes place in vitro by inseminating the oocyte with processed spermatozoa. After fertilization, one or more embryo(s) are transferred into the uterine cavity. Indications for IVF are tubal pathology, unexplained subfertility, endometriosis, mild male subfertility and failure to conceive after ovarian stimulation or after intrauterine insemination (http://www.nvog-documenten.nl/uploaded/docs/09_indicaties_ivf.pdf).

The controlled ovarian stimulation technique used in IVF is designed to induce multifollicular growth in order to maximize the yield of oocytes for fertilization to enhance the probability of achieving a pregnancy. Oocyte retrieval is performed by ultrasound guided transvaginal follicle aspiration. The oocytes are inseminated in vitro with processed semen after which the embryos are cultured for a variable number of days before the embryos will be transferred transcervical to the uterine cavity. Surplus transferable embryos will be cryopreserved.

Intracytoplasmatic sperm injection (ICSI)
Intracytoplasmatic sperm injection is an assisted reproductive technique in which

oocytes are retrieved and become mechanically injected in vitro with one single spermatozoon per oocyte. After this, fertilization takes place in vitro and the embryo is transferred into the uterine cavity. The indications for ICSI treatment are severe male factor or no fertilization after IVF procedure. Couples undergoing the ICSI treatment have similar procedures for ovarian stimulation, oocyte retrieval and embryo transfer as IVF treatment. The difference between ICSI and IVF is the oocyte fertilization procedure. Both treatments are in vitro but with ICSI, the oocytes are injected with one spermatozoon, while with IVF the oocytes are inseminated with many thousands of spermatozoa.

1.3.2 Risks and complications of ART

Numerous studies have shown that ART is an efficacious treatment in subfertile couples. In spite of its established efficacy, little attention has been paid so far to the safety of ART, i.e. to its adverse events and complications to the female patient and the offspring.

Complications to the female patient

The adverse effects of ART for the female patient are mainly ovarian hyperstimulation syndrome (OHSS) and multiple pregnancy.

- Ovarian hyperstimulation syndrome (OHSS)

The most serious and potentially lethal -iatrogenic- complication of ovarian stimulation is ovarian hyperstimulation syndrome (OHSS)[58]. OHSS is a self-limiting syndrome and ranges from mild to severe. The incidence of the severe form of OHSS in patients undergoing ART is about 1% but this is increasing due to the expansion of ART for the last years[59].

- Multiple pregnancy

One of the most important complications of ART treatments is the high percentage of multiple pregnancies. The number of twin births increased rapidly since the mid-1970s in the Netherlands from 10/1000 births to 18/1000 births in the year 2000, but this is decreasing to 16/1000 births in 2005 (www.cbs.nl). Also the number of triplets and higher order gestations increased from the mid-1970s (0.15/1000 births) to 1994 (0.6/1000 births), but since then this number is decreasing to 0.2/1000 births in 2005 (www.cbs.nl).

Increase in delayed childbearing and increasing use of assisted reproductive treatments (like IVF) are the most important contributors to the almost 2-fold increase in twin birth between 1975 and 2000. Older women are at increased risk for

multiple gestations in naturally conceived pregnancies. They are also at increased risk of being infertile and of using ART, placing them at higher risk of multiple gestations. The increase of triplets and higher order gestations between 1975 and 1995 was partly due to the high numbers of embryos (four or five) transferred after IVF/ICSI. In the early 1990s this number was reduced to two or three embryos and resulted in mainly a reduction of triplets and higher order births. Possibly due to the application of single embryo transfers in IVF/ICSI, a decrease in twin birth rate is notified since 2004[60]. Multiple pregnancies are nowadays considered as the most important complication resulting from ART and are related with maternal and perinatal mortality and morbidity.

- Maternal mortality

The causes of maternal death in developed countries are mainly due to hypertensive disorders, thromboembolism and haemorrhage[61]. Advanced maternal age increases the risk of maternal mortality rate in IVF (and other related assisted reproductive technologies) pregnancies.

- Maternal morbidity

One of the major maternal complications associated with multiple pregnancies are higher incidence of pregnancy-related hypertensive conditions, like eclampsia and preeclampsia. This risk is two times greater for a twin pregnancy than for a singleton pregnancy, and a 3-fold greater risk for a triplet pregnancy[62].

Maternal morbidity is also significantly increased in multiple compared to singleton pregnancies due to pregnancy complications like venous thromboembolism, antepartum haemorrhage from placenta praevia and placenta abruption, anaemia, gestational diabetes, delivery by Caesarean section, preterm delivery and postpartum haemorrhage[63-66].

- Long-term effects of ART on women

The long-term effects of hormonal ovarian stimulation on the risk of breast cancer and cancers of the female genital tract have been investigated in cohort studies where infertile women exposed to ovarian stimulation treatment were compared with infertile women without ovarian stimulation[67-70]. These studies showed no increased risk for cancer of the breast or ovary in women who received hormonal ovarian stimulation and those who did not. Also a lower incidence of cancer of the breast and carcinoma in situ of the cervix was found in women with live birth following IVF treatment compared with women with live birth without IVF[71]. The average follow-up time in most studies was relatively short, further studies are necessary to include longer follow-up times to reveal any possible long-term effects of ART on women.

Complications for offspring

Perinatal mortality rate for multiple pregnancy is higher compared to singleton pregnancy. This rate increases with the number of gestation[72] due to a higher rate of prematurity and low birth weights in children. The perinatal outcome of ART pregnancies shows that singleton ART pregnancies have a slightly increased perinatal morbidity and mortality compared to naturally conceived singleton pregnancies [13-16]. Although perinatal outcomes are similar for twins conceived by ART and naturally conceived twins[14,17,22-25] the twin pregnancy itself is the main course of the overall poorer perinatal outcome of twins.

Many perinatal complications can be addressed to prematurity and low birth weight. Therefore longer treatment is required in neonatal intensive care unit for twins compared to singletons with more health costs[73].

- Long term consequences on offspring

Long term consequences on neurlogical sequelae in IVF/ICSI pregnancies is still unclear. Two population based studies showed that twins born after assisted conception have a similar risk of neurological sequelae (cerebral palsy) as naturally conceived twins and singletons after IVF or ICSI[74,75]. A recent review showed no increased risk of imprinting diseases in children born after use of assisted reproductive technologies compared with children born after naturally conception[76].

- Health of male offspring

Concern has also been expressed about the health of boys born to couples with male factor subfertility. Genetic causes of male factor subfertility include: congenital bilateral absence of the vas deference (CBAVD)/cystic fibrosis, Y-chromosomal microdeletions, X-chromosomal and autosomal aberrations (i.e. Robertsonian translocations), syndromal disorders featuring infertility (i.e. Kallmann´s syndrome) and ultrastructural sperm defects with a genetic basis. It is theoretically possible that with ICSI these genetic abnormalities may be transmitted to the following male generation[77]. However, more large prospective studies are needed to address this theory.

The incidence of de-novo sex chromosomal aberrations is higher after ICSI than in the general population[78]. Also the major malformation rate (eg hypospadias) after ICSI might be higher as compared with natural conception[79]. This may be due to the infertility *per se,* rather than the ART technique. Despite concern about the health of the male offspring, there is yet no strong evidence for ICSI procedure related risks[80].

1.4 Outline of the thesis

In this thesis two subjects of investigation on evaluation of diagnosis and therapy in reproductive medicine are presented. The evaluation of diagnosis in reproductive medicine is based on a study on hysterosalpingography (HSG) (chapters 2 to 5) and describes the results of a pragmatic multicenter randomized controlled trial comparing fertility work-ups with or without HSG to assess the value of HSG prior to laparoscopy and dye in a routine clinical setting. The question whether HSG should be maintained as a tool in the basic fertility work-up is of major interest as otherwise many subfertile women will continue to undergo a diagnostic procedure with questionable utility. The second subject of this thesis focusses on therapy in reproductive medicine and is based on a study on assisted conception (chapter 6) to clarify whether adverse perinatal outcome is related to assisted conception or other maternal risk factors, and whether this is for both singleton and twin pregnancies.

In chapter **two**, the effect of HSG on cumulative pregnancy rate is assessed. We performed a randomized trial to investigate the influence of routine use of HSG at an early stage in the fertility work-up prior to laparoscopy and dye on cumulative pregnancy rate, compared with routine use of laparoscopy without HSG. In chapter **three** we evaluate pain and patient satisfaction of women undergoing an HSG or laparoscopy and dye as routine diagnostic procedures in the fertility work-up in a cohort study of subfertile women nested in the randomized controlled trial. Chapter **four** describes the diagnostic properties of HSG and Chlamydia antibody testing. Nested in the randomized controlled trial, the value of Chlamydia antibody testing is assessed as alternative test for HSG. In chapter **five** the results of recruitment difficulties in a randomized trial involving HSG are described. Chapter **six** describes the results of a systematic review of controlled studies to examine whether there are genuine differences in outcome between assisted and natural conceptions and whether they apply to both singleton and twin pregnancies. Finally, in chapter **seven** the results of all studies are provided and its consequences for future approaches are discussed.

References

1. Cochrane AL. 1931-1971: a critical review, with particular reference to the medical profession. In: Medicines for the year 2000. London: Office of Health Economics, 1979, 1-11.
2. Chalmers I, editor (1988 – 1992). *Oxford Database of Perinatal Trials*. Oxford: Oxford University Press. (Version 1.0, Disk Issue 1).
3. Chalmers I, Enkin M, Keirse MJNC. Effective care in pregnancy and childbirth. Oxford: Oxford University Press, 1989.
4. Enkin M, Keirse MJNC, Chalmers I. A guide to effective care in pregnancy and childbirth. Oxford: Oxford University Press, 1989.
5. Helmerhorst FM, Belfield T, Kulier R, Maitra N, O'Brien P, Grimes DA. The Cochrane Fertility Regulation Group: Synthesizing the best evidence about family planning. Contraception 2006;74:280-286.
6. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995;273:408–412.
7. Schulz KF, Chalmers I. Grimes DA, Altman DG. Assessing the quality of randomisation from reports of controlled trials published in obstetrics and gynecology journals. JAMA 1994;272:125-128.
8. Haynes B. Can it work? Does it work? Is it worth it? The testing of healthcare interventions is evolving. BMJ 1999;319:652-653.
9. Vandenbroucke JP. Balancing benefits and harms in health care: observational data on harm should complement systematic reviews of benefit. BMJ 2003;327:750.
10. Glasziou P, Vandenbroucke JP, Chalmers I. Assessing the quality of research. BMJ 2004;328:39-41.
11. Writing Group for the Women´s Health Initiative Investigators. Risk and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women. Principal Results From the Women´s Health Initiative Randomized Controlled Trial. JAMA 2002;288:321-333.
12. Steptoe PC, Edwards RG. Birth after the reimplantation of a human embryo. Lancet 1978;2:366.
13. Australian In Vitro Fertilisation Collaborative Group. High incidence of preterm births and early losses in pregnancy after in vitro fertilisation. BMJ 1985;291:1160-1163.
14. Dhont M, De Sutter P, Ruyssinck G, Martens G, Bekaert A. Perinatal outcome of pregnancies after assisted reproduction: a case-control study. Am J Obstet Gynecol 1999;181:688-695.
15. Tan SL, Doyle P, Campbell S, Beral V, Rizk B, Brinsden P, et al. Obstetric outcome of in vitro fertilization pregnancies compared with normally conceived pregnancies. Am J Obstet Gynecol 1992;167:778-784.
16. Koudstaal J, Braat DD, Bruinse HW, Naaktgeboren N, Vermeiden JP, Visser GH. Obstetric outcome of singleton pregnancies after IVF: a matched control study in four Dutch university hospitals. Hum Reprod 2000;15:1819-1825.
17. Dhont M, De Neubourg F, Van der Elst J, De Sutter P. Perinatal outcome of pregnancies after assisted reproduction: a case-control study. J Assist Reprod Genet 1997:14:575-580.
18. Isaksson R, Gissler M, Tiitinen A. Obstetric outcome among women with unexplained infertility after IVF: a matched case-control study. Hum Reprod 2002;17:1755-1761.
19. Moise J, Laor A, Armon Y, Gur I, Gale R. The outcome of twin pregnancies after IVF. Hum Reprod

1998;13:1702-1705.

20. Tallo CP, Vohr B, Oh W, Rubin LP, Seifer DB, Haning RV Jr. Maternal and neonatal morbidity associated with in vitro fertilization. J Pediatr 1995;127:794-800.
21. Koudstaal J, Bruinse HW, Helmerhorst FM, Vermeiden JP, Willemsen WN, Visser GH. Obstetric outcome of twin pregnanceis after in-vitro fertilization: a matched control study in four Dutch university hospitals. Hum Reprod 2000;15:935-940.
22. Agustsson T, Geirsson RT, Mires G. Obstetric outcome of natural and assisted conception twin pregnancies is similar. Acta Obstet Gynecol Scand 1997;76:45-49.
23. Minakami H, Sayama M, Honma Y, Matsubara S, Koike T, Sato I, Uchida A, Eguchi Y, Momoi M, Araki S. Lower risks of adverse outcome in twins conceived by artificial reproductive techniques compared with spontaneously conceived twins. Hum Reprod 1998;13:2005-2008.
24. Fitzsimmons BP, Bebbington MW, Fluker MR. Perinatal and neonatal outcomes in multiple gestations: assisted reproduction versus spontaneous conception. Am J Obstet Gynecol 1998;179:1162-1167.
25. Olivennes F, Kadhel P, Rufat P, Fanchin R, Fernandez H, Frydman R. Perinatal outcome of twin pregnancies obtained after in vitro fertilization: a comparison with twin pregnancies obtained spontaneously or after ovarian stimulation. Fertil Steril 1996;66:105-109.
26. de Graaf A. Onderzoek gezinsvorming 2003. Voorburg/Heerlen: Centraal Bureau voor de Statistiek, 2004.
27. National Institute for Clinical Excellence (NICE) 2004. Fertility: Assessment and treatment for people with fertility problems-full guideline. London, RCOG Press. Website: http://www.rcog.org.uk/resources/Public/pdf/Fertility_full.pdf
28. Jansen FW, Kapiteijn K, Trimbos-Kemper T, Hermans J, Trimbos JB. Complications of laparoscopy: a prospective multicenter observational study. Br J Obstet Gynaecol 1997;104:595-600.
29. Rzymski P, Wosniak J, Opala T, Wilczak M, Sajdak S. Anaphylactic reaction to methylene blue dye after laparoscopic chromopertubation. Int J Gynaecol Obstet 2003;81:71-72.
30. World Health Organization 1999. WHO laboratory manual for the examination of human semen and semen-cervical mucus interaction, 4th ed., Cambridge University Press.
31. Mol BW, Collins JA, Burrows EA, van der Veen F, Bossuyt PM. Comparison of hysterosalpingography and laparoscopy in predicting fertility outcome. Hum Reprod 1999;14:1237-1242.
32. Royal College of Obstetricians and Gynaecologists 1998. The initial investigation and management of the infertile couple. Evidence-Based Clinical Guidelines No 2, RCOG Press, London, UK.
33. Al-Fadhli R, Sylvestre C, Buckett W, Tulandi T. A randomized study of laparoscopic chromopertubation with lipiodol versus saline in infertile women. Fertil Steril 2006;85:505-507.
34. Rindfleisch W. Darstellung des cavum uteri. Berl klin Wschr 1910;47:780-781.
35. Rubin IC. Nonoperative determination of patency of fallopian tubes in sterility. Intrauterine inflation with oxygen, and production of an artificial penumoperitoneum. Preliminary report. J Am Med Assoc 1920;74:1017.
36. Heuser C. Lipiodol in the diagnosis of pregnancy. Lancet 1925;2:1111-1112.
37. Schuitemaker NW, Helmerhorst FM, Tjon a Tham RT, van Saase JL. Late anaphylactic shock after

hysterosalpingography. Fertil Steril 1990;54:535-536.

38. Pittaway DE, Winfield AC, Maxson W, Daniell J, Herbert C, Wentz AC. Prevention of acute pelvic inflammatory disease after hysterosalpingography: efficacy of doxycycline prophylaxis. Am J Obstet Gynecol 1983;147:623-626.
39. Stumpf PG, March CM. Febrile morbidity following hysterosalpingography: identification of risk factors and recommendations for prophylaxis. Fertil Steril 1980;33:487-492.
40. Soules MR, Spadoni LR. Oil versus aqueous media for hysterosalpingography: a continuing debate based on many opinions and few facts. Fertil Steril 1982;38:1-11.
41. Lindequist S, Justesen P, Larsen C, Rasmussen F. Diagnostic quality and complications of hysterosalpingography: oil- versus water-soluble contrast media-a randomized prospective study. Radiology 1991;179:69-74.
42. Johnson N, Vandekerckhove P, Watson A, Lilford R, Harada T, Hughes E. Tubal flushing for subfertility. Cochrane Database Syst Rev 2005;2:CD003718.
43. Glatstein IZ, Harlow BL, Homstein MD. Practice patterns among reproductive endocrinologists: further aspects of the infertility evaluation. Fertil Steril 1998;70:263-269.
44. Spring DB, Barkan HE, Pruyn SC. Potential therapeutic effects of contrast materials in hysterosalpingography: a prospective randomized clinical trial. Kaiser Permanente Infertility Work Group. Radiology. 2000 ;214:53-57.
45. Yun AJ, Lee PY. Enhanced fertility after diagnostic hysterosalpingography using oil-based contrast agents may be attributable to immunomodulation. Am J Roentgenol 2004;183:1725-1727.
46. Shalev J, Meizner I, Bar-Hava I, Dicker D, Mashiach R, Ben-Rafael Z. Predictive value of transvaginal sonography performed before routine diagnostic hysteroscopy for evaluation of infertility. Fertil Steril 2000;73:412-417.
47. Soares SR, Barbosa dos Reis MM, Camargos AF. Diagnostic accuracy of sonohysterography, transvaginal sonography, and hysterosalpingography in patients with uterine cavity diseases. Fertil Steril 2000;73:406-411.
48. Swart P, Mol BW, van der Veen F, van Beurden M, Redekop WK and Bossuyt PM. The accuracy of hysterosalpingography in the diagnosis of tubal pathology: a meta-analysis. Fertil Steril1995;64:486-491.
49. Mol BW, Swart P, Bossuyt PM, van Beurden M, van der Veen F. Reproducibility of the interpretation of hysterosalpingography in the diagnosis of tubal pathology. Hum Reprod 1996;11:1204-1208.
50. Glatstein IZ, Sleeper LA, Lavy Y, Simon A, Adoni A, Palti Z, Hurwitz A, Laufer N. Observer variability in the diagnosis and management of the hysterosalpingogram. Fertil Steril 1997;67:233-237.
51. Mol BW, Swart P, Bossuyt PM, van der Veen F. Is hysterosalpingography an important tool in predicting fertility outcome? Fertil Steril 1997;67:663-669.
52. Maas JWM, Evers JLH, ter Riet G, Kessels AGH. Pregnancy rate following normal versus abnormal hysterosalpingography findings: a meta-analysis. Gynecol Obstet Invest 1997;43:79-83.
53. te Velde ER, Boer-Meisel ME, Meisner J, Schoemaker J and Habbema JD. The significance of preoperative hysterosalpingography and laparoscopy for predicting the pregnancy outcome in patients with a bilateral hydrosalpinx. Eur J Obstet Gynecol Reprod Biol 1989;31:33-45.

54. Peipert JF. Clinical practice. Genital chlamydial infections. N Engl J Med 2003;349:2424-2430.
55. Punnonen R, Terho P, Nikkanen V and Meurman O. Chlamydial serology in infertile women by immunofluorescence. Fertil Steril 1979;31:656–659.
56. Den Hartog JE, Morre SA, Land JA. Chlamydia trachomatis-associated tubal factor subfertility: Immunogenetic aspects and serological screening. Hum Reprod Update 2006;12:719-730.
57. Mol BW, Collins JA, Van Der Veen F and Bossuyt PM. Cost-effectiveness of hysterosalpingography, laparoscopy, and Chlamydia antibody testing in subfertile couples. Fertil Steril 2001;75:571-580.
58. The Practice Committee of the American Society for Reproductive Medicine. Ovarian hyperstimulation syndrome. Fertil Steril 2006;86:S178-S183.
59. Delvigne A, Rozenberg S. Epidemiology and prevention of ovarian hyperstimulation syndrome (OHSS): a review. Hum Reprod Update 2002;8:559-577.
60. Vilska S, Tiitinen A, Hyden-Granskog C, Hovatta O. Elective transfer of one embryo results in an acceptable pregnancy rate and eliminates the risk of multiple birth. Hum Reprod 1999;14:2392-2395.
61. Khan KS, Wojdyla D, Say L, Gülmezoglu AM, Van Look PF. WHO analysis of causes of maternal death: a systematic review. Lancet 2006;367:1066-1074.
62. Day MC, Barton JR, O'Brien JM, Istwan NB, Sibai BM. The effect of fetal number on the development of hypertensive conditions of pregnancy. Obstet Gynecol 2005;106:927-931.
63. Sebire NJ, Jolly M, Harris J, Nicolaides KH. Risks of obstetric complications in multiple pregnancies: an analysis of more than 400 000 pregnancies in the UK. Prenat Neonat Med 2001;6:89-94.
64. Wen SW, Demissie K, Yang Q, Walker MC. Maternal morbidity and obstetrical complications in triplet pregnancies and quadruplet and higher-order multiple pregnancies. Am J Obstet Gynecol 2004;191:254-258.
65. Walker MC, Murphy KE, Pan S, Yang Q, Wem SW. Adverse maternal outcomes in multiple pregnancies. BJOG 2004;111:1294-1296.
66. Buhling KJ, Henrich W, Starr E, Lubke M, Bertram S, Siebert G, Dudenhausen JW. Risk for gestational diabetes and hypertension for women with twin pregnancy compared to singleton pregnancy. Arch Gynecol Obstet 2003;269:33-36.
67. Venn A, Watson L, Lumley J, Giles G, King C, Healy D. Breast and ovarian cancer incidence after infertility and in vitro fertilisation. Lancet 1995;346:995-1000.
68. Venn A, Watson L, Bruinsma F, Giles G, Healy D. Risk of cancer after use of fertility drugs with in-vitro fertilisation. Lancet 1999;354:1586-1590.
69. Dor J, Lerner-Geval L, Rabinovici J, Chetrit A, Levran D, Lunenfeld B, Mashiach S, Modan B. Cancer incidence in a cohort of infertile women who underwent in vitro fertilization. Fertil Steril 2002;77:324-327.
70. Doyle P, Maconochie N, Beral V, Swerdlow AJ, Tan SL. Cancer incidence following treatment for infertility at a clinic in the UK. Hum Reprod 2002;17**:**2209–2213.
71. Kristiansson P, Björ O, Wramsby H. Tumour incidence in Swedish women who gave birth following IVF treatment. Hum Reprod 2007;22:421-426.
72. Doyle P. The outcome of multiple pregnancy. Hum Reprod 1996;11S4:110-117.

73. Lukassen HG, Schonbeck Y, Adang EM, Braat DD, Zielhuis GA, Kremer JA. Cost analysis of singleton versus twin pregnancies after in vitro fertilization. Fertil Steril. 2004;81:1240-1246.
74. Lidegaard O, Pinborg A, Andersen AN. Imprinting diseases and IVF: Danish National IVF cohort study. Hum Reprod 2005;20:950-954.
75. Pinborg A, Loft A, Schmidt L, Greisen G, Rasmussen S, Andersen AN. Neurological sequelae in twins born after assisted conception: controlled national cohort study. BMJ 2004;329:311.
76. Lidegaard O, Pinborg A, Andersen AN. Imprinting disorders after assisted reproductive technologies. Curr Opin Obstet Gynecol 2006;18:293-296.
77. Meschede D, Lemcke B, Behre HM, De Geyter C, Nieschlag E, Horst J. Clustering of male infertility in the families of couples treated with intracytoplasmic sperm injection. Hum Reprod 2000;15:1604-1608.
78. Bonduelle M, Van Assche E, Joris H, Keymolen K, Devroey P, Van Steirteghem A, Liebaers I. Prenatal testing in ICSI pregnancies: incidence of chromosomal anomalies in 1586 karyotypes and relation to sperm parameters. Hum Reprod 2002;17:2600-2614.
79. Ericson A, Kallen B. Congenital malformations in infants born after IVF: a population-based study. Hum Reprod 2001;16:504-509.
80. Devroey P, Van Steirteghem A. A review of ten years experience of ICSI. Hum Reprod Update 2004;10:19-28.

Chapter

2

Routine use of hysterosalpingography prior to laparoscopy in the fertility workup: a multicenter randomised controlled trial

D.A.M.Perquin[1], P.J.Dörr[1], A.J.M.de Craen[2] and F.M.Helmerhorst[3]

[1] Department of Obstetrics and Gynaecology, Medical Center Haaglanden, The Hague, The Netherlands

[2] Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands

[3] Department of Gynaecology, Division of Reproductive Medicine, Leiden University Medical Center, Leiden, The Netherlands

Hum Reprod 2006; 21 : 1227-31

Abstract

***Background*:** A multicenter randomised controlled trial with or without hysterosalpingography (HSG) was conducted to assess the usefulness of HSG as a routine investigation in the fertility workup prior to laparoscopy and dye.

Methods: From 1 April 1997 to 1 April 2002, subfertile women were allocated by a computer–based 1 : 1 ratio randomisation procedure, either for an HSG followed by laparoscopy and dye (the intervention group) of for laparoscopy and dye only (the control group) as a part of their fertility workup. Cumulative pregnancy rate (CPR) within 18 months after randomisation was the primary outcome of interest.

Results: 344 women were randomised to the intervention group (n = 169) and the control group (n = 175). There was no significant difference in CPR at 18 months in the intervention group (49.1%) [95% confidence interval (CI) 41.6 to 56.6] and the control group (50.3%) (95% CI 42.8 to 57.8), a difference of –1.2% (95% CI –11.8% to 9.5%).

***Conclusion*:** The routine use of HSG at an early stage in the fertility workup prior to laparoscopy and dye does not influence CPR, compared with the routine use of laparoscopy and dye without HSG.

Introduction

After history taking, physical examination, semen analysis and investigation of ovulation, assessment of tubal patency is the next step in the standard examination of the subfertile couple. Owing to the noninvasive nature and low cost, hysterosalpingography (HSG) is widely used as a first-line approach to assess the patency of the Fallopian tubes in routine fertility workup[1,2], although laparoscopy and dye is considered the gold standard[3,4].

A reason for performing HSG instead of or prior to laparoscopy and dye cannot be found in the test characteristics of HSG. Comparing the accuracy of HSG with that of laparoscopy and dye in the diagnosis of tubal pathology, a meta-analysis demonstrated point estimates of 65% of sensitivity and 83% of specificity[4]. Furthermore, considerable variability in the interpretation as well as clinical consequences of HSG abnormalities has been shown among practitioners[5,6]. Advantages of HSG relative to laparoscopy are the short outpatient procedure and the enhancement of pregnancy with oil-soluble contrast medium[7], although water-soluble media are mostly used[8]. The therapeutic effect of tubal flushing with water-soluble media is, however, still unknown[9].

The relative merits of HSG and laparoscopy in screening for tubal factors have been discussed for more than 30 years, but so far no randomised controlled trial has been reported[1]. To assess the value of HSG prior to laparoscopy and dye in a routine clinical setting, we performed a pragmatic multicenter randomised controlled trial comparing fertility workups with or without HSG. In a pragmatic trial, effectiveness of an intervention is assessed under usual circumstances, in contrast to efficacy trials in which the intervention is examined under ideal conditions[10]. Is the patient better off with or without the extra intervention (in this case, HSG)? We compared the two strategies, with pregnancy as a clinical endpoint, in terms of cumulative pregnancy rate (CPR).

Subjects and methods

Patients and randomisation procedure

The study was performed in three teaching hospitals in The Netherlands. All newly referred and admitted subfertile women who visited the Department of Reproductive Medicine of Leiden University Medical Center (April 1997 to April 2002), the Department of Obstetrics and Gynaecology of the Medical Center Haaglanden, The Hague (April 1997 to April 2002) or the Department of Obstetrics and Gynaecology of the Groene Hart Hospital, Gouda, The Netherlands (April 1999 to April 2000)

were eligible for inclusion in the trial.
Exclusion criteria were subfertility less than 1 year, woman older than 37 years at the time of first visit, anovulation despite clomiphene citrate or bromocriptine use, abnormal semen analysis according to World Health Organization (WHO)[11] criteria or testing of tubal patency performed in the past. The institutional review boards of each of the three hospitals approved the study protocol. Women were asked to participate in the study by their treating gynaecologist at the time that HSG would normally be planned, and informed consent was obtained. The treating gynaecologist telephoned the secretariat of Medical Center Haaglanden at The Hague to perform randomisation. A computer-based 1 : 1 ratio randomisation procedure was used to allocate the women into two groups. Randomisation was stratified for each participating hospital. All women routinely received vaginal ultrasound before randomisation. The intervention group underwent HSG first, and if the HSG showed normal uterine cavity and no tubal pathology and if the woman did not conceive within 6 months, a laparoscopy and dye followed after 6 months. When tubal pathology was assumed, laparoscopy was performed within 1–2 months after the HSG. The control group received a laparoscopy and dye immediately. If pathology of the uterine cavity was presumed by HSG or by vaginal ultrasound, hysteroscopy could be performed together with the laparoscopy. Moreover, a history of recurrent miscarriages or diethylstilboestrol (DES) exposure was an additional reason to perform a hysteroscopy during laparoscopy.
Because our trial was designed to determine the effectiveness of HSG in the routine fertility workup, we ensured that HSG and laparoscopy results were uniformly interpreted in all participating hospitals. At the same time, the study protocol intentionally allowed normal clinical freedom and a variety of choices and protocols after HSG and laparoscopy. Hence, the participating hospitals used their own protocol for therapeutic reproductive surgery and assisted reproductive treatments [e.g. intrauterine insemination (IUI) or IVF]. The primary analysis was conducted on an intention-to-treat basis. The primary outcome parameter in our study was occurrence of pregnancy within 18 months after randomisation. The diagnosis of pregnancy was based on a positive urine or serum pregnancy test in association with the presence of an intrauterine gestation sac on ultrasound scan.

HSG and laparoscopy and dye

All hysterosalpingographies were performed in the outpatient clinic of the department of radiology shortly after the menstrual period. A water-soluble contrast medium (Omnipaque 300®) was used. One photograph was taken of the phase when the cavity and tubes were just filled and one when there was overflow at both sides or when there was maximal filling of the tubes without overflow. After 30 min, a late

film was made to detect contrast depots. Findings of tubal pathology at HSG were classified according to[2], as normal, one-sided abnormality or two-sided abnormality. Additional intracavity abnormalities were scored separately. The results of HSG were interpreted in a weekly meeting by staff members specialized in reproductive medicine, who also decided whether laparoscopy and dye should be performed with or without delay.

Laparoscopy and dye was performed in the follicular phase and under general anaesthesia. After making pneumoperitoneum, a thorough inspection of the pelvis, internal genitalia, appendix and liver region was performed, followed by testing the patency of the Fallopian tubes using dye. A dilute solution of Methylene Blue was injected through the cervix. During laparoscopy, we determined adhesions, structural abnormalities of the uterus, endometriosis, periadnexal disease and Fallopian tube occlusion. Tubal pathology at laparoscopy was defined according to Mol *et al*.[2], as normal, one-sided abnormality or two-sided abnormality. Furthermore, endometriosis detected at laparoscopy was classified according to the classification of the American Fertility Society[12]. Therapeutic reproductive surgery could be applied during laparoscopy, such as coagulation of endometriosis grade I/II, laparoscopic adhesiolysis or laparoscopic cystectomy.

Statistical methods

Descriptive statistics were used to assess the similarity of the groups. Categorical data were assessed by the chi-square test and continuous variables by Student's *t*-test. CPRs were calculated using standard time-to-event analysis (Kaplan–Meier survival analysis). For comparison of the different CPR curves, the log-rank statistic was used. On the basis of local unpublished data of Leiden University Medical Center, we calculated that for a subfertile couple the probability of getting pregnant after 1 year from intake, including artificial interference, is about 45%. With a smallest difference in CPR arbitrarily set at 10% (55% in the intervention group and 45% in the control group), an alpha error of 0.05 and a beta error of 0.20 (power of the study set at 80%), we calculated that at least 375 women should be included in each arm (a total of 750 women).

Results

A total of 344 women were randomised, 169 to the intervention group and 175 to the control group. Follow-up either to pregnancy or for 18 months was complete for all subjects in both groups. Figure 1 shows the flow chart of participants. At the end of the study, HSG had been performed in 152 of the 169 (90%) women in the

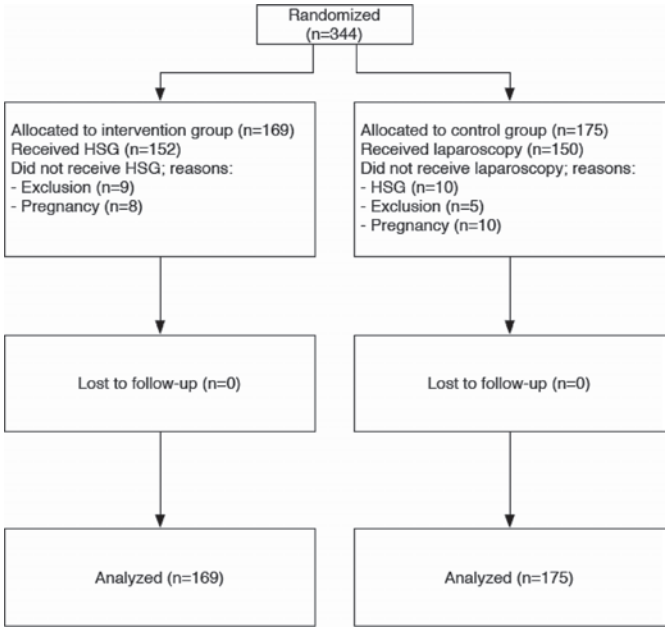


Figure 1.
Flow chart of participants.

intervention group. In the control group, 10 of the 175 (6%) women had undergone an HSG. Laparoscopies had been performed on 94 of the 169 (56%) women in the intervention group and on 150 of the 175 (86%) women in the control group. To deal with this, our analysis was based on the groups as randomised, following the intention-totreat principle.

Table I summarizes the baseline characteristics of the women participating in the study. When comparing the women in the intervention and control groups, there were no significant differences between the groups regarding age, duration of subfertility, parity, history of pelvic inflammatory disease, intrauterine device, sexually transmitted disease or tubal surgery.

A total of 152 women in the intervention group underwent HSG, and in 58 women no additional laparoscopy was performed. In 40 of these 58 women, this was due to occurrence of pregnancy within 6 months after HSG. Hence, the pregnancy rate within 6 months after performing HSG was 26%. Of the other 18 women, no laparoscopy was performed for unknown reasons.

At HSG, 23 (15%) showed one-sided abnormalities and 23 (15%) showed two-sided abnormalities. In 10 of 152 women, HSG showed intracavity abnormalities, a reason why hysteroscopy was performed during laparoscopy.

Ninety-four laparoscopies were performed in the intervention group. At laparoscopy, 12 (13%) showed one-sided abnormalities and 15 (16%) showed two-sided abnormalities. Endometriosis was detected in 21 women (22%). The laparoscopic

Table I. Baseline characteristics of women participating in the study

Characteristics	Intervention group (n = 169)	Control group (n = 175)
Age (years) (mean ± SD)	30 ± 4.0	30 ± 4.0
Duration of subfertility (years) (mean ± SD)	2.0 ± 1.1	2.0 ± 0.8
Nulliparous	138 (82)	145 (83)
History of		
Pelvic inflammatory disease	5 (3)	5 (3)
Intrauterine device	4 (2)	5 (3)
Sexually transmitted disease	6 (4)	7 (4)
Tubal surgery	2 (1)	1 (1)
Ovulatory disorder	31 (18)	25 (14)

Values are expressed as number (percentage) unless otherwise specified.

Table II. Therapeutic reproductive surgery applied in the intervention and control groups

Surgery	Intervention group (n = 94)	Control group (n = 150)
None	65 (69)	102 (68)
Coagulation of endometriosis grade I/II	23 (25)	41 (27)
Laparoscopic adhesiolysis	5 (5)	5 (4)
Laparoscopic cystectomy	1 (1)	2 (1)

Values are expressed as number (percentage).

therapeutic interventions applied in the intervention group are summarized in Table II.

In 150 of 175 women in the control group, a laparoscopy and dye was performed. Seven of these 150 women also had a hysteroscopy performed during laparoscopy. At laparoscopy, 13 (9%) showed one-sided abnormalities and 16 (11%) showed two-sided abnormalities. Endometriosis was detected in 43 women (29%). Different laparoscopic therapeutic interventions could be applied when abnormal findings were seen (Table II). Table II shows that there were no significant differences in numbers of laparoscopic therapeutic interventions between both groups.

Table III shows the numbers of assisted reproductive treatments applied during the study in the intervention group and in the control group. There were no significant differences in numbers of applied assisted reproductive treatments between both groups.

In the intervention group, 83 pregnancies occurred in follow-up time of 2115 months, whereas in the control group 88 pregnancies occurred in 2257 months of follow-up. Therefore, the annualized pregnancy rate was 0.47 [95% confidence interval (CI)

Table III. Assisted reproductive treatment applied in the intervention and control groups

Treatment	Intervention group (n = 169)	Control group (n = 175)
None	94 (56)	101 (58)
Intrauterine insemination	60 (35)	61 (35)
IVF	15 (9)	13 (7)

Values are expressed as number (percentage).

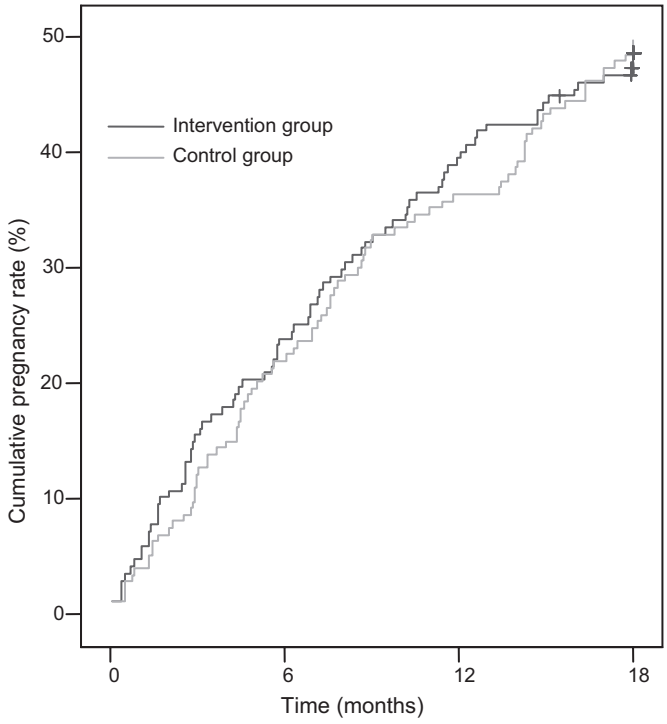


Figure 2. Cumulative pregnancy rates for the intervention group, which included hysterosalpingography (HSG), and control group, which excluded HSG.

0.37 to 0.57] in the intervention group and 0.47 (95% CI 0.37 to 0.57) in the control group. The CPR at 18 months was 49.1% (95% CI 41.6 to 56.6) in the intervention group and 50.3% (95% CI 42.8 to 57.8) in the control group, a difference of –1.2% (95% CI –11.8 to 9.5). Pregnancy rates at given times throughout the study were also not significantly different (Figure 2). No difference in primary or secondary subfertility was found. Of the 31 multiparous women in the intervention group, 15

were pregnant at the end of follow-up, compared with 17 of 30 multiparous women in the control group. Moreover, 69% of women in the intervention group who received assisted reproductive treatment were pregnant at the end of follow-up, compared with 63% in the control group. The 6% difference between study groups was not statistically significant (95% CI –8 to 21).

Directly after randomisation, it appeared that 14 couples were inappropriately randomised. The reasons were as follows: abnormal semen analysis (*n*=9), maternal age (*n*=2) and immediate withdrawal of women from participation in the trial but they agreed to be followed up for pregnancy (*n*=3). In the remaining 330 women, the CPR at 18 months was 50% (95% CI 42 to 58) in 160 women in the intervention group and 51% (95% CI 43 to 59) in 170 women in the control group. This difference of –1% was not significant (95% CI –12 to 10).

When excluding all women who did not receive the allocated intervention, 17 in the intervention group and 25 in the control group, the CPR at 18 months was 46% (95% CI 38 to 54) in the 152 women in the intervention group and 45% (95% CI 37 to 53) in the 150 women in the control group.

Discussion

Here, we report results from a randomised controlled trial assessing the usefulness of HSG as a conventional diagnostic investigation in the subfertility workup prior to laparoscopy, the gold standard for detecting tubal pathology. There was no difference in CPRs at 18 months between the intervention group (49.1%) and the control group (50.3%). We demonstrated that routine use of HSG prior to laparoscopy and dye in the fertility workup did not result in a significant effect on the incidence of pregnancy, compared with routine use of laparoscopy without HSG.

The distribution of laparoscopic therapeutic interventions and additional assisted reproductive treatments was equal between the intervention and control groups (Tables II and III). Hence, it is unlikely that these laparoscopic therapeutic interventions and additional assisted reproductive treatments affected the difference between both groups. Where early laparoscopy is not routine, the trial results also show that a protocol that begins with HSG and proceeds to laparoscopy when indicated has a similar yield to that of early routine laparoscopy.

A flaw of this study is the limited number of women who entered the trial. Our original power calculation envisaged 375 women in each group. We ended enrolment of women after 5 years because of difficulties in recruitment, thereby accepting a lower power for the study. By that time, 344 women were participating, each group containing about 200 less than the original power calculations required. Given the

absolute equality of a close to 50% pregnancy rate in both groups, a pregnancy rate of about 65% would have been required in the next 200 women in the intervention group to achieve overall significance at the 5% level. The chances of this happening were very small, given a success rate of about 50% in the first 169 women. As the binomial probability of such an occurrence is less than 1 in 100, it is extremely unlikely that continuing the study would have altered the conclusions that can be drawn from it.

The additional value of HSG compared with laparoscopy in the fertility workup is particularly the assessment of the uterine cavity. Uterine cavity malformations with a frequency of about 10–15% in subfertile women can be visualized by HSG, although the effectiveness of surgical treatment at enhancing pregnancy rates is not established[9]. The endovaginal ultrasound has been demonstrated as a reliable alternative for detecting uterine cavity pathology[13]. It can also be used in the visualization of pelvic pathology such as endometriosis or ovarian pathology. Ultrasonography in comparison with HSG is less invasive and easily performed in a short period without roentgen radiation. HSG, though, remains the only trustworthy method for examining tubal mucosal fold configuration, especially when tubal microsurgery (e.g. salpingostomy) is considered[14], albeit that HSG then can be better carried out after laparoscopy. The reason that HSG is performed early in the fertility workup is based more on tradition and personal preference, rather than on the demonstrated usefulness of its components[1]. In these circumstances, oil-soluble contrast medium can be used to improve pregnancy rates, although HSG with oil-soluble contrast medium is associated with an increased risk of granuloma formation and embolism; mortality in connection with embolism has been reported[15].

Chlamydia antibody testing (CAT) may replace HSG as screening test that can be used for estimating the risk of tubal pathology prior to laparoscopy. Several studies have shown that CAT might be as accurate as HSG in diagnosing tubal pathology [16-19].

In conclusion, this study shows that routine use of HSG at an early stage in the fertility workup prior to laparoscopy and dye does not influence cumulative pregnancy rate.

Acknowledgements

We thank Erik Hallensleben, MD, Groene Hart Hospital, Gouda, for his contribution to the trial. Additional thanks to Miranda de Vreede and Marian Mackay-Bosch, Medical Center Haaglanden, The Hague, for logistical assistance.

References

1. Helmerhorst FM, Oei SG, Bloemenkamp KW and Keirse MJHC (1995) Consistency and variation in fertility investigations in Europe. Hum Reprod 10,2027–2030.
2. Mol BW, Collins JA, Van Der Veen F and Bossuyt PM (2001) Cost-effectiveness of hysterosalpingography, laparoscopy, and Chlamydia antibody testing in subfertile couples. Fertil Steril 75,571–580.
3. Rowe PJ, Comhaire FH, Hargreave TB and Mellows HJ (1993) *WHO Manual for the Standardized Investigation and the Diagnosis of the Infertile Couple*. Cambridge University Press, Cambridge, UK, p. 4.
4. Swart P, Mol BW, van der Veen F, van Beurden M, Redekop WK and Bossuyt PM (1995) The accuracy of hysterosalpingography in the diagnosis of tubal pathology: a meta-analysis. Fertil Steril 64,486–491.
5. Mol BW, Swart P, Bossuyt PM, van Beurden M and van der Veen F (1996) Reproducibility of the interpretation of hysterosalpingography in the diagnosis of tubal pathology. Hum Reprod 11,1204–1208.
6. Glatstein IZ, Sleeper LA, Lavy Y, Simon A, Adoni A, Palti Z, Hurwitz A and Laufer N (1997) Observer variability in the diagnosis and management of the hysterosalpingogram. Fertil Steril 67,233–237.
7. Johnson N, Vandekerckhove P, Watson A, Lilford R, Harada T and Hughes E (2005) Tubal flushing for subfertility. Cochrane Database Syst Rev 18,CD003718.
8. Glatstein IZ, Harlow BL and Hornstein MD (1998) Practice patterns among reproductive endocrinologists: further aspects of the infertility evaluation. Fertil Steril 70,263–269.
9. National Institute for Clinical Excellence (2004) *Fertility: Assessment and Treatment for People with Fertility Problems—Full Guideline*. RCOG Press, London.
10 Haynes B (1999) Can it work? Does it work? Is it worth it? The testing of healthcare interventions is evolving. BMJ 319,652–653.
11. World Health Organization (1999) *WHO Laboratory Manual for the Examination of Human Semen and Semen–Cervical Mucus Interaction*, 4th edn. Cambridge University Press, Cambridge, UK.
12. American Fertility Society (1985) Revised American Fertility Society classification of endometriosis: 1985. Fertil Steril 43,351–352.
13. Braun P, Grau FV, Pons RM and Enguix DP (2005) Is hysterosalpingography able to diagnose all uterine malformations correctly? A retrospective study. Eur J Radiol 53,274–279.
14. te Velde ER, Boer-Meisel ME, Meisner J, Schoemaker J and Habbema JD (1989) The significance of preoperative hysterosalpingography and laparoscopy for predicting the pregnancy outcome in patients with a bilateral hydrosalpinx. Eur J Obstet Gynecol Reprod Biol 31,33–45.
15. Lindequist S, Justesen P, Larsen C and Rasmussen F (1991) Diagnostic quality and complications of hysterosalpingography: oil-versus water-soluble contrast media – a randomised prospective study. Radiology 179,69–74.

16. Dabekausen YA, Evers JL, Land JA and Stals FS (1994) Chlamydia trachomatis antibody testing is more accurate than hysterosalpingography in predicting tubal factor infertility. Fertil Steril 61,833–837.
17. Mol BW, Dijkman B, Wertheim P, Lijmer J, van der Veen F and Bossuyt PM (1997) The accuracy of serum chlamydial antibodies in the diagnosis of tubal pathology: a meta-analysis. Fertil Steril 67,1031–1037.
18. Land JA, Evers JL and Goossens VJ (1998) How to use Chlamydia antibody testing in subfertility patients. Hum Reprod 13,1094–1098.
19. Veenemans LM and van der Linden PJ (2002) The value of Chlamydia trachomatis antibody testing in predicting tubal factor infertility. Hum Reprod 17,695–698.

Chapter

3

Evaluation of pain and patient satisfaction of women undergoing hysterosalpingography or laparoscopy and dye

Denise A.M. Perquin, MD,[1] P. Joep Dörr, MD, PhD,[1]
Anton J.M. de Craen, PhD[2] and Frans M. Helmerhorst, MD, PhD[3]

[1] Department of Obstetrics and Gynaecology, Medical Center Haaglanden, The Hague, The Netherlands

[2] Department of Gerontology and Geriatrics, Leiden University Medical Center, The Netherlands

[3] Department of Gynaecology, Division of Reproductive Medicine, Leiden University Medical Center, The Netherlands

(submitted for publication)

Abstract

Objectives: To evaluate pain and patient satisfaction of women undergoing hysterosalpingography (HSG) or laparoscopy and dye as routine diagnostic procedures in the fertility work-up.

Study design: One hundred and thirty-six subfertile women were randomized to HSG or laparoscopy. Pain was assessed before the procedure and again after the procedure at time of discharge by using a visual analogue scale (VAS) and the McGill Pain Questionnaire-Dutch Language Version (MPQ-DLV). One week after the procedure, physical complaints, analgesic usage, and overall judgment of the procedure was assessed by a postal questionnaire.

Results: Shortly after the procedure, women who underwent an HSG (n=73) reported the procedure to be significantly more painful compared to women who underwent laparoscopy (n=50). Women who underwent laparoscopy had significantly more complaints (shoulder pain, sore throat, and muscular pain), required more analgesics, and were unable to return to normal daily activities for a longer interval of time. Results of the overall judgment of the procedure favored laparoscopy over HSG.

Conclusions: The HSG procedure is experienced as more painful while laparoscopy and dye is related with more post-procedure complaints. The overall judgment of women undergoing the procedures favors laparoscopy.

Introduction

Hysterosalpingography (HSG) and laparoscopy and dye are two widely used diagnostic interventions in the fertility work-up. Due to the non-invasive, outpatient procedure and low cost, HSG is usually applied as first-line approach to assess tubal pathology[1,2]. However, according to the World Health Organization (WHO) criteria, laparoscopy and dye is generally considered the gold standard[3]. Almost 80% of U.S. board-certified reproductive endocrinologists routinely perform HSG and laparoscopy[4].

HSG has been criticized for its moderate sensitivity and specificity in the diagnosis of tubal factor subfertility compared to laparoscopy and dye[5]. Moreover, one of the drawbacks of the HSG procedure is the painful nature of the intervention[6]. We prospectively evaluated pain and patient satisfaction of women undergoing an HSG or laparoscopy and dye as diagnostic intervention in the fertility work-up. In a group of women who underwent either HSG or laparoscopy and dye we assessed pain after the intervention, and evaluated physical complaints, use of analgesics, and overall judgment of the procedure one week after the procedure.

Material and methods

Subjects

All women in our study participated in a multicenter randomized controlled trial to assess the usefulness of HSG as routine investigation in the fertility work-up prior to laparoscopy and dye. Recruitment strategy, description of subjects, and main results of the study have been published elsewhere[7]. In short, the trial was performed in three teaching hospitals in the Netherlands. All newly referred and admitted subfertile women who visited one of the three hospitals between April 1997 and April 2002 were eligible for inclusion. Exclusion criteria were subfertility less than 1 year, age of woman older than 37 years, anovulation in spite of clomifene citrate or bromocriptin use, abnormal semen analysis according to WHO criteria[8], or testing of tubal patency performed in the past. A total number of 344 women were randomized, 169 to the HSG group and 175 to the laparoscopy group. The HSG group received an HSG, which was followed by a laparoscopy and dye after six months unless pregnancy occurred. The laparoscopy group received a laparoscopy and dye immediately. For the present analysis we used data of the first intervention, which was either the HSG or the laparoscopy. Laparoscopies performed after the HSG were not included in the analysis. All subjects gave informed consent. The

Institutional Review Boards of each of the three hospitals approved all stages of the study.

HSG and laparoscopy and dye

Women were asked to participate in the trial by their treating gynaecologist at the time that HSG would normally be planned and counseled them about performing an HSG or laparoscopy. A detailed standardized patient information sheet on both procedures was given to all women when the procedure was planned. All women who underwent HSG received a prescription for three tablets of naproxennatrium 275 mg and were instructed to take two tablets 30-60 minutes before HSG and one tablet directly after HSG. A senior resident in gynaecology or a consultant specialized in reproductive medicine performed HSG at the department of Radiology. A suction Semm cup (A, B or C) was placed over the cervix and suction was applied via a vacuum pump. In less than 5% of all HSGs adequate suction could not be maintained. In that case a tenaculum was applied by grasping the anterior lip of the cervix and a Foley balloon catheter or metallic Whitehead cannula was applied to align the cervical canal. A water-soluble contrast medium (Omnipaque 300; Nycomed Ltd, Birmingham, UK) was used. Dye was injected slowly to potentially minimize discomfort. The amount of contrast used was at most 10 ml and depended on the time necessary to assess tubal status.

Laparoscopy and dye was performed by a senior resident in gynaecology supervised by a consultant specialized in reproductive medicine. All women had general anaesthesia, the majority was induced and maintained with propofol and remifentanyl. No preoperative analgesics or sedatives were given. After induction of general anaesthesia, a double puncture (peri-umbilical and suprapubic) technique was used with a 10-mm 0 degree non-disposable laparoscope and one 5 mm trocar. A balloon-tipped catheter was placed inside the cervical canal for testing tubal patency with methylene blue dye. A pneumoperitoneum was created by insufflating carbon dioxide. After completion of the procedure, carbon dioxide was allowed to fully escape. Pelvic findings were recorded on a standardized datasheet. All women were given standardized treatment for postoperative pain (paracetamol and nonsteroidial anti-inflammatory drugs). Anti-emetic drugs were not given routinely.

Pain scores and questionnaire

Subjects rated their overall pain by using a visual analogue scale (VAS) and the McGill Pain Questionnaire, Dutch language version (MPQ-DLV). The VAS was drawn as a 10-cm line, anchored by word descriptions at zero cm “no pain” and at 10 cm “worst pain imaginable”. It had 1 cm major tick marks, and 1 mm minor tick

marks. At each major tick mark the corresponding number was printed.

The McGill Pain Questionnaire (MPQ)[9] is a standardized questionnaire and has been translated into Dutch (McGill Pain Questionnaire, Dutch language version (MPQ-DLV))[10]. This translated questionnaire is considered to be highly reliable[10]. The lists of pain descriptors were given to the patients who were asked to choose one word in each subscale (sensory, items 1 through 12; affective, items 13 through 17; evaluative, item 18 through 20) that best described their pain. The frequency with which each descriptor was chosen by the patients and the total number of words chosen (NWC) were assessed. Subscale pain rating index (PRI) was calculated by summing item scores within a subscale. A total PRI was also calculated by summing all subscale scores.

The patients were asked to fill in the VAS and the MPQ-DLV before the procedure and again after the procedure at time of discharge. The first pain score ("pre-procedure score") was measured before the procedure to obtain a baseline value. This pain score consisted of the mean and maximal abdominal pain experienced during the first day of the last menstrual period. The second pain score ("post-procedure score") was assessed after the procedure at time of discharge. On this occasion, both pain experienced during the procedure and pain at the time of assessment were measured. The patient was asked to rate the mean and maximum abdominal pain during the procedure itself and their pain at the moment of completion the pain scores. All subjects themselves recorded the pain scores on the VAS and the MPQ-DLV.

One week after the procedure a questionnaire was sent by post addressing issues like assessment of physical complaints during the last week (10 items), analgesic usage, and overall judgment of the procedure (3 statements). The three statements assessing the overall judgment of women who experienced either HSG or laparoscopy were as follows: procedure turned out badly, no repeat of procedure under same circumstances if required, and advice against procedure to a well known friend. All complaints were rated on a five-point scale of none, mild, moderate, severe, or very severe.

Statistical methods

Descriptive statistics were used to assess similarity of intervention groups. Categorical data were assessed by Chi-square tests and continuous variables by Student's t-test. Responses to the pain scores and questionnaire were assessed with paired t-tests. We dichotomized the rating of complaints by combining none and

mild outcomes into one category and moderate, severe, and very severe in an other category. P-values <0.05 were considered significant.

Results

We assessed pain and patient satisfaction in a convenience sample of 123 women. There were no differences between the women who were included and women who were not included in the present analysis, except for parity (89% of included women were nulliparous *vs* 79% of women who were not included, p=0.03) and subfertility (80% of included women had primary subfertility *vs* 68% of women who were not included, p=0.02). No significant differences in baseline characteristics were found between the 73 women in the HSG group and the 50 women in the laparoscopy group (table 1).
The pre- and post procedure VAS pain scores and current pain intensity for the HSG and laparoscopy group are shown in table 2. As expected, there were no differences in the mean and maximal pre-procedure pain scores between both study groups. After the intervention, women who underwent an HSG reported the procedure to be more painful compared to women who underwent a laparoscopy and dye (mean VAS pain score, 46.7 *vs* 35.7 mm; p=0.004 and maximal VAS pain score, 66.4 *vs* 51.0 mm; p<0.001). However, the pain experienced at the time of the assessment (current pain intensity) was significantly lower in the HSG group than in the laparoscopy group (23.4 *vs.* 31.3 mm; p=0.016).

Table 3 shows the increase in the score of the sensory, affective, and evaluative dimensions of the number of words chosen and the pain-rating index of the MPQ-DLV for the HSG and laparoscopy group. All scores on the pain dimensions in the HSG group were positive, indicating a worsening of pain experience. All scores on the pain dimensions in the laparoscopy group were negative. Five of the eight comparisons between the two groups yielded a p-value <0.1, of which one was <0.05.

Table 1. Baseline characteristics of the 123 women who were included in the study.

Characteristics	HSG (n = 73)	Laparoscopy (n = 50)
Age (years), mean ± SD	29.8 ± 4.0	30.2 ± 3.4
Duration of subfertility (years), mean ± SD	2.0 ± 1.4	2.0 ± 0.9
Nulliparous	61 (84)	45 (90)
Primary subfertility	56 (77)	41 (82)

Values are expressed as number (percentage) unless otherwise specified.

Table 2. VAS pain scores pre- and post procedure for HSG and laparoscopy group.

Pain score	HSG (n = 73)	Laparoscopy (n = 50)	P-value
Mean pain (SE)			
Pre-procedure	30.7 (2.3)	32.0 (2.7)	0.72
Post-procedure	46.7 (2.8)	35.7 (2.2)	0.004
Increase	16.0 (3.2)	3.7 (2.8)	0.006
Maximal pain (SE)			
Pre-procedure	41.5 (3.0)	45.8 (3.1)	0.33
Post-procedure	66.4 (2.8)	51.0 (2.9)	<0.001
Increase	24.9 (3.4)	5.2 (3.6)	<0.001
Current pain intensity (SE)	23.4 (2.3)	31.3 (2.2)	0.016

SE = Std. Error
Values are expressed as number (percentage) unless otherwise specified.

Pre-procedure is the measurement, before the procedure, of abdominal pain experienced during the first day of the last menstrual period; Post-procedure means, after the procedure, measurement of overall pain related to the procedure.

Table 3. Difference in pre- and post procedure painscores for HSG and laparoscopy groups measured by MPQ-DLV.

	HSG (n = 73)	Laparoscopy (n = 50)	P-value
NWC (SE)			
NWC-s	0.35 (0.31)	-0.15 (0.38)	0.31
NWC-a	0.04 (0.15)	-0.54 (0.20)	0.02*
NWC-e	0.08 (0.11)	-0.15 (0.12)	0.16
NWC-t	0.47 (0.48)	-0.84 (0.61)	0.09
PRI (SE)			
PRI-s	0.58 (0.55)	-1.00 (0.65)	0.07
PRI-a	0.15 (0.23)	-0.59 (0.33)	0.06
PRI-e	0.25 (0.30)	-0.25 (0.32)	0.26
PRI-t	0.99 (0.92)	-1.80 (1.20)	0.06

MPQ-DLV= McGill Pain Questionnaire – Dutch Language Version
NWC-s. –a, -e, -t = Numbers of words chosen-sensory, -affective, -evaluative, and total scale
PRI -s. –a, -e, -t = Pain rating index-sensory, -affective, -evaluative, and total scale

SE = standard error

* = Significant difference between two groups

Table 4. One week evaluation of HSG and laparoscopy.

	HSG (n = 73)	Laparoscopy (n = 50)	P-value
Complaints			
Abdominal pain	36 (49)	23 (46)	0.72
Vaginal blood loss	13 (18)	14 (28)	0.18
Shoulder pain	7 (10)	34 (68)	<0.001
Nausea	5 (7)	6 (21)	0.33
Vomiting	2 (3)	2 (4)	0.70
Dizziness	5(7)	9 (18)	0.06
Headache	13 (18)	7 (14)	0.57
Sore throat	7 (10)	21 (42)	<0.001
Muscular pain	8 (11)	19 (38)	<0.001
Tiredness	27 (37)	23 (46)	0.32
Use of pain medication	15 (21)	24 (48)	<0.001
Mean no. of days unable to perform daily activities (± SE)	1.3 (0.3)	3.5 (0.2)	<0.001
Patient satisfaction			
Procedure turned out badly	26 (36)	7 (14)	<0.01
No repeat of procedure under same circumstances if required	17 (23)	3 (6)	0.01
Advise against procedure to a well known friend	14 (19)	3 (6)	0.04

SE = standard error
Values are expressed as number (percentage) unless otherwise specified

Table 4 shows the physical complaints experienced during the first week after the performance of HSG or laparoscopy. Almost 50% of women in both groups had abdominal pain after the intervention. The most frequently reported complaints in the HSG group were abdominal pain (49%) and tiredness (37%), and in the laparoscopy group shoulder pain (68%), abdominal pain (46%), and tiredness (46%). Significantly more complaints were mentioned after laparoscopy and dye than HSG for shoulder pain ($p<0.01$), sore throat ($p<0.01$), and muscular pain ($p<0.01$). Women in the laparoscopy group needed more analgesics ($p<0.001$) and were unable to return to normal daily activities for a longer interval of time ($p<0.001$). All three statements regarding patient satisfaction were statistically significant in favor of the laparoscopy.

Comment

We evaluated pain and patient satisfaction of women undergoing an HSG or laparoscopy and dye as routine diagnostic investigation in the fertility work-up. After the procedure, HSG was experienced as a more painful procedure than laparoscopy. One week after the procedure women who underwent laparoscopy had significantly more complaints (shoulder pain, sore throat, and muscular pain), required more analgesics, and were unable to return to normal daily activities for a longer interval of time. Patient satisfaction indicated that women favored laparoscopy over HSG.

We assessed, under usual circumstances, the experience of HSG or laparoscopy and dye in an effectiveness trial. In an effectiveness trial the intervention is assessed under usual circumstances. This is in contrast with an efficacy trial in which the intervention is examined under ideal conditions[11]. Pragmatism was also the reason that the post procedure pain score was assessed after a shorter time in the HSG group relative to the laparoscopy group. The time of pain assessment was usually 30 min after the HSG procedure, whereas women in the laparoscopy group filled in their pain scores at time of discharge usually 4-5 hours after the procedure when the effect of most narcotics has worn off.

The overall mean and maximal visual analogue scale (VAS) pain scores were significantly higher in the HSG group, whereas pain experienced at the moment of VAS completion was significantly higher in the laparoscopy group. This observation can be explained by the peak tissue reaction caused by the abrupt injection of the contrast medium during HSG that dies out shortly thereafter[12].
Reasons of pain experienced during HSG are related with the phase of the procedure: cervical instrumentation, uterine distension with contrast media, and peritoneal irritation a result of contrast spill into the peritoneal cavity[13]. The cervical cup cannula appears to be superior in decreasing discomfort compared to the traditional metal cannula[14]. Second choice was the use of the Foley balloon catheter or the metal cannula. We performed all HSGs with water-soluble contrast medium, which is common practice[15]. Two randomized studies have demonstrated no difference in pain experience between oil- and water-soluble contrast media[16,17]. The presence of tubal blockage on HSG, which may cause severe pain, was also documented. This had no significant effect on pain scores, which is in agreement with other studies[12,16,18].

HSG and laparoscopy are experienced as highly stressful procedures and cause physical complaints in more than half of women[6,19]. Age, duration, and type of

subfertility appear to have no significant effect on physical complaints. Adequate counselling may reduce fear and may increase pain tolerance[20] e.g. by focusing on realistical positive expectations[21]. Although effective analgesia during HSG is important, two prospective studies showed no difference in pain experience with or without intrauterine topical local anesthetic at HSG[12,22].

Shoulder pain (65%) was the most mentioned complaint one week after laparoscopy. This finding, which is in line with other studies[21], is thought to be due to residual gas irritating the peritoneum under the diaphragm, where it can remain for up to 3 days. Women in the laparoscopy group needed also significantly more analgesics than women in the HSG group and were more days unable to perform daily activities than women in the HSG group, that might also related to pneumoperitoneum. If post-operative nausea is to be avoided in ambulatory gynecological patients, it may be necessary to decrease or eliminate the use of narcotics in anesthetic induction. Epidural anesthesia may be an alternative to general anesthesia in gynecologic laparoscopy to reduce abdominal pain and better postoperative recovery, although the pain during manipulation might hinder the procedure[24].

The one-week evaluation indicated that women who underwent HSG had less physical complaints than women who underwent laparoscopy and dye. Eimers *et al.*[6] came to opposite conclusions. That study was about physical complaints and emotional stress related to routine diagnostic procedures in the fertility work-up and was not randomized and described a heterogeneous group.

Compared with laparoscopy, HSG is less invasive with no risk to anesthetic and surgical complications. The additional value of HSG is the assessment of the uterine cavity but the effectiveness of surgical treatment to enhance pregnancy rates has not been established[25]. If tubal pathology has been detected and microsurgery is considered, HSG is the only method to evaluate mucosal fold configuration. One of the main disadvantages of HSG is that it has moderate diagnostic accuracy for tubal patency compared to laparoscopy[5].
Laparoscopy and dye is the reference standard for assessing tubal status and enables direct inspection of the pelvis including endometriosis and pelvic adhesions. Additionally, laparoscopic treatments such as diathermy of endometriosis or adhesiolysis may be applied at the same time. The risk of having peritubal abnormalities or endometriosis in subfertile women who are asymptomatic, have normal physical examination and bilateral tubal patency at HSG is low. On these issues laparoscopy could be delayed or abandoned because in a minority of women severe tubal pathology will be revealed[25-27].

In conclusion, HSG is experienced as more painful than laparoscopy and dye, however laparoscopy is related with more post-procedure complaints. The overall judgment of women who experienced either HSG or laparoscopy is in favor of laparoscopy and dye.

Acknowledgements

The authors would like to thank Anne M. Stiggelbout, MD, PhD, from the Department of Medical Decision Making, Leiden University Medical Center, for making the questionnaire.

References

1. Helmerhorst FM, Oei SG, Bloemenkamp KW, Keirse MJHC. Consistency and variation in fertility investigations in Europe. Hum Reprod 1995; 10: 2027-30.
2. Mol BW, Collins JA, van der Veen F, Bossuyt PM. Cost-effectiveness of hysterosalpingography, laparoscopy, and Chlamydia antibody testing in subfertile couples. Fertil Steril 2001; 75: 571-80.
3. Rowe PJ, Comhaire FH, Hargreave TB and Mellows HJ. WHO manual for the standardized investigation and the diagnosis of the infertile couple. Cambridge University Press, 1993.
4. Glatstein IZ, Harlow BL, Hornstein MD. Practice patterns among reproductive endocrinologists: the infertility evaluation. Fertil Steril 1997; 67: 443-51.
5. Swart P, Mol BW, van der Veen F, van Beurden M, Redekop WK, Bossuyt PM. The accuracy of hysterosalpingography in the diagnosis of tubal pathology: a meta-analysis. Fertil Steril 1995; 64: 486-91.
6. Eimers JM, Omtzigt AM, Vogelzang ET, van Ommen R, Habbema JDF, te Velde ER. Physical complaints and emotional stress related to routine diagnostic procedures of the fertility investigation. J Psychosom Obstet Gynaecol 1997; 18: 31-5.
7. Perquin DAM, Dörr PJ, de Craen AJM, Helmerhorst FM. The routine use of hysterosalpingography prior to laparoscopy in the fertility work-up: a multicenter randomized controlled trial. Hum Reprod 2006; 21: 1227-31.
8. World Health Organization. WHO laboratory manual for the examination of human semen and semen-cervical mucus interaction, 4th ed. Cambridge: Cambridge University Press, 1999.
9. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. Pain 1975; 1: 277-99.
10. van der Kloot WA. Oostendorp RAB, van der Meij J, van den Heuvel J. The Dutch version of the McGill pain questionnaire: a reliable pain questionnaire. Ned Tijdschr Geneeskd 1995; 139: 669-73.

11. Haynes B. Can it work? Does it work? Is it worth it? The testing of healthcare interventions is evolving BMJ 1999; 319: 652–3.
12. Costello MF, Horrowitz S, Steigrad S, Saif N, Bennett M, Ekangaki A. Transcervical intrauterine topical local anesthetic at hysterosalpingography: a prospective, randomized, double-blind, placebo-controlled trial. Fertil Steril 2002; 78: 1116-22.
13. Moore DE. Pain associated with hysterosalpingography: Ethiodol versus Salpix media. Fertil Steril 1982; 38: 629-31.
14. Cohen SB, Wattiez A, Seidman DS *et al*. Comparison of cervical vacuum cup cannula with metal cannula for hysterosalpingography. BJOG 2001; 108: 1031-5.
15. Glatstein IZ, Harlow BL, Hornstein MD. Practice patterns among reproductive endocrinologists: further aspects of the infertility evaluation. Fertil Steril 1998; 70: 263-9.
16. Lindequist S, Justesen P, Larsen C, Rasmussen F. Diagnostic quality and complications of hysterosalpingography: oil-versus water-soluble contrast media--a randomized prospective study. Radiology 1991; 179: 69-74.
17. Lindequist S, Rasmussen F, Larsen C. Use of iotrolan versus ethiodized poppy-seed oil in hysterosalpingography. Radiology 1994; 191: 513-7.
18. Sohail S. Variables affecting immediate pain tolerance in X-ray hysterosalpingography. J Coll Physicians Surg Pak 2004; 14: 170-2.
19. Kopitzke EJ, Berg BJ, Wilson JF, Owens D. Physical and emotional stress associated with components of the infertility investigation: perspectives of professionals and patients. Fertil Steril 1991; 55: 1137-43.
20. Weller A, Hener T. Invasiveness of medical procedures and state anxiety in women. Behav Med 1993; 19: 60-5.
21. Koyama T, McHaffie JG, Laurienti PJ, Coghill RC. The subjective experience of pain: where expectations become reality. Proc Natl Acad Sci U S A 2005; 102: 12950-5.
22. Frishman GN, Spencer PK, Weitzen S, Plosker S, Shafi F. The use of intrauterine lidocaine to minimize pain during hysterosalpingography: a randomized trial. Obstet Gynecol 2004; 103: 1261-6.
23. Codd C. Are analgesics necessary for women at home following laparoscopic gynaecological day surgery? Nurs Pract 1991; 5: 8-12.
24. Kuramochi K, Osuga Y, Yano T *et al.* Usefulness of epidural anesthesia in gynecologic laparoscopic surgery for infertility in comparison to general anesthesia. Surg Endosc 2004; 18: 847-51.
25. National Institute for Clinical Excellence. Fertility: assessment and treatment for people with fertility problems. National Collaborating Center for Women´s and Children´s Health. London: RCOG Press, 2004.
26. Royal College of Obstetricians and Gynaecologists. The initial investigation and management of the infertile couple. Evidence-Based Clinical Guidelines No 2. London: RCOG Press, 1998.
27. Erel CT, Senturk LM. Is laparoscopy necessary before assisted reproductive technology? Curr Opin Obstet Gynecol 2005; 17: 243-8.

Chapter

4

The value of Chlamydia trachomatis–specific IgG antibody testing and hysterosalpingography for predicting tubal pathology and occurrence of pregnancy

Denise A. M. Perquin, M.D.[1], Matthias F. C. Beersma, M.D., Ph.D.[2], Anton J. M. de Craen, Ph.D.[3], Frans M. Helmerhorst, M.D., Ph.D.[4]

[1] Department of Obstetrics and Gynaecology, Medical Center Haaglanden, The Hague, The Netherlands

[2] Department of Virology, Erasmus Medical Center, Rotterdam, The Netherlands

[3] Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands

[4] Division of Reproductive Medicine, Department of Gynecology, Leiden University Medical Center, Leiden, The Netherlands

Fertil Steril 2007; 88 (1): 224-6

Abstract

We assessed two diagnostic methods, *Chlamydia trachomatis*–specific IgG and hysterosalpingography (HSG), as a screening test for the likelihood of tubal damage or occurrence of pregnancy before laparoscopy in 178 subfertile women who were randomly assigned to HSG followed by laparoscopy or immediate laparoscopy. The diagnostic accuracy and prognostic value of both *C. trachomatis*–specific IgG antibody testing and HSG are comparable but show poor performance.

Introduction

Tubal pathology accounts for approximately 14% of the causes of subfertility[1]. Therefore, assessment of the tubal status has become a routine part of the fertility work-up. Several diagnostic tests can be used to assess tubal status, of which *Chlamydia* antibody testing (CAT) and hysterosalpingography (HSG) comprise the first-line approach. These tests are usually followed by the gold standard laparoscopy and dye[2]. The CAT and HSG tests provide risk estimates of tubal pathology before laparoscopy, but the diagnosis of tubal pathology can only be made with laparoscopy and dye. The diagnostic accuracy of CAT and HSG compared with laparoscopy and dye is well established[3–7]. Most studies have shown that CAT performs just as well or better than HSG. However, the prognostic value of both tests in predicting occurrence of pregnancy is not well known. Idahl *et al*.[8] showed that there were no differences in achieving pregnancy between *C. trachomatis* IgG negative or positive women. Moreover, as shown by several studies, the prognostic value of HSG also is low[9–13].

Material and methods

The aim of the present study was twofold. First, we assessed the diagnostic accuracy of *C. trachomatis*–specific IgG (CtsIgG) and HSG compared with laparoscopy and dye in a large group of subfertile women. Second, we used both tests as prognostic indicator to assess occurrence of pregnancy.

All women in this study participated in a multicenter, randomized controlled trial with or without the performance of HSG to assess the usefulness of HSG as routine investigation in the fertility work-up before laparoscopy and dye. Recruitment strategy, description of subjects, and the main results of this study have been

published elsewhere[14]. The present part of the trial took place only at the Division of Reproductive Medicine, Department of Gynecology, Leiden University Medical Center. The institutional review board of the hospital approved all stages of the study.

As part of the fertility work-up, blood was drawn at the patient's first visit. All spare serum was cryopreserved. For the present study, the spare serum of the participating women was thawed to perform a species-specific peptidebased serologic assay for the presence of IgG antibodies to *C. trachomatis* (SeroCT; Savyon Diagnostics, Ashdod, Israel). The results of CtsIgG antibody testing were compared with the findings of HSG in the diagnosis of tubal pathology using laparoscopy and dye as the reference standard. The results of both tests were also compared with occurrence of pregnancy within 18 months after randomisation as the clinical end point in terms of cumulative pregnancy rate. The diagnosis of clinical ongoing pregnancy was based on positive urine or serum pregnancy test in association with an intact intrauterine gestation sac on ultrasound scans.

The HSG was performed with a water-soluble contrast medium (Omnipaque 300, Nycomed Ltd, Birmingham, UK). An HSG was considered abnormal if there was occlusion of one or both tubes or peritubal adhesions without tubal occlusion. Laparoscopy and tubal testing was performed with methylene blue dye. Laparoscopic findings of tubal factor subfertility were defined as extensive periadnexal adhesions and/or distal occlusion of one or both tubes[15]. Endometriosis was not included.

The SeroCT IgG assay and calculation were performed according to the manufacturer's instructions. The qualitative outcome of the assay (cut-off index) was categorized as negative (< 1.0) or positive (> 1.10). Values in the equivocal zone (optical densities between the values for negativity and positivity) were considered as negative results in the calculations when a second sample was also borderline or negative.

The diagnostic accuracy of CtsIgG and HSG compared with laparoscopy were assessed by calculating sensitivity, specificity, and likelihood ratios. The prognostic value of CtsIgG and HSG in predicting occurrence of pregnancies was assessed with Kaplan-Meier survival analysis and Cox proportional hazard models. *P* values of .05 were considered to be statistically significant.

Results

A total of 178 subfertile women participated in the present study. Sixty-five women underwent both HSG and laparoscopy, 25 women only HSG, and 88 women only laparoscopy. Therefore, the CtsIgG results could be compared with laparoscopy in 153 women, whereas the HSG results could be compared with laparoscopy in 65

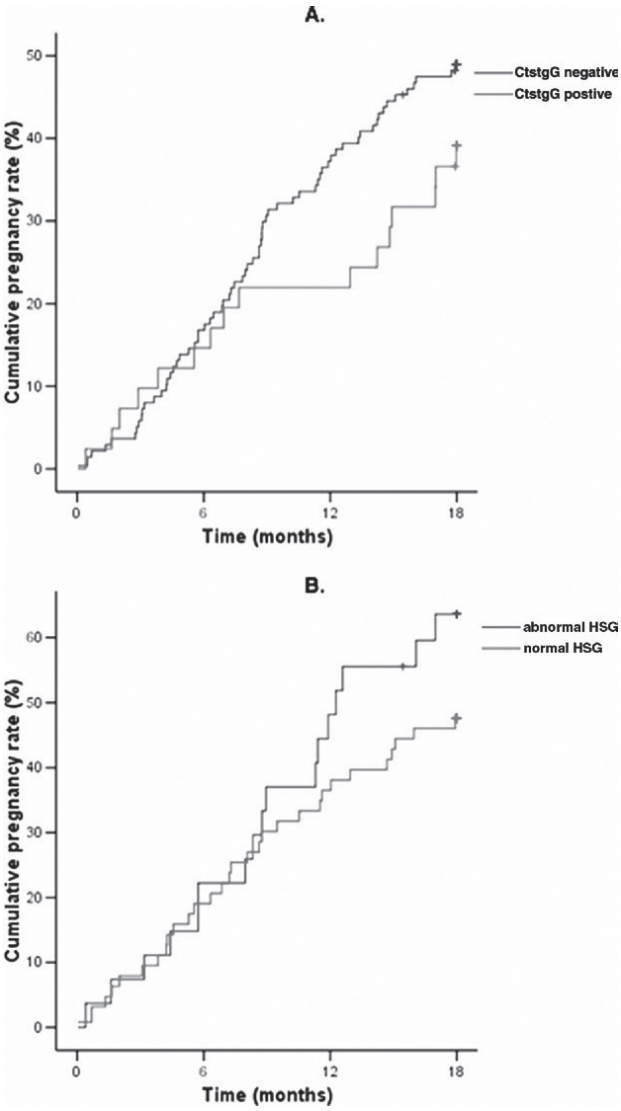


Figure 1. Cumulative pregnancy rates between CtslgG result and HSG findings.

women. Out of all 153 women, 9 (6%) showed unilateral occlusion, 14 (9%) showed bilateral occlusion, and 8 (5%) showed peritubal adhesions without tubal occlusion at laparoscopy.

In 35 out of 153 women studied (23%), CtsIgG was positive; 14 of these seropositive women (40%) had tubal pathology according to laparoscopy. Out of the 118 CtsIgG-negative women, 17 had tubal pathology (14%). The sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio for CtsIgG in predicting tubal pathology were 45%, 83%, 2.6, and 0.7, respectively.

In 65 out of 153 women, an HSG and laparoscopy were performed. A total of 41 women had a normal HSG, and 24 had an abnormal HSG. Fifteen women (23%) showed unilateral occlusion, eight (12%) showed bilateral occlusion, and one (2%) showed peritubal adhesions without tubal occlusion. In 11 out of 24 women with abnormal HSG (46%), tubal pathology was found at laparoscopy. Out of the 41 women with normal HSG, 5 had tubal pathology at laparoscopy (12%). The sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio for HSG were 69%, 73%, 2.6, and 0.4, respectively.

For women who achieved pregnancy during the trial period, whether spontaneous or treatment related, there were no statistically significant differences found in cumulative pregnancy rate between CtsIgG-negative and CtsIgG-positive results (hazard ratio 0.73, 95% confidence interval [CI] 0.42–1.25; *P* .25) (Fig. 1A). Pregnancy rates at given times throughout the study were also not significantly different. There was also no significant difference in cumulative pregnancy rate between normal and abnormal findings at HSG (hazard ratio 1.33, 95% CI 0.73–2.41; *P* .35) (Fig. 1B).

Discussion

The main strength of the present study is the prospective design. All women included had a diagnostic laparoscopy for evaluation of tubal status. An allocation concealment was applied with 0% loss to follow-up. The shortcoming of the study is that the diagnostic or prognostic value of CAT and HSG can be determined separately, but the total sample size is too small for the combination of diagnostic and prognostic value.

We showed that *C. trachomatis* IgG antibodies have comparable diagnostic accuracy with HSG in detecting tubal pathology. These findings have been shown

by other studies that also compared CAT with HSG and used laparoscopy as reference standard[3–7]. Those authors' measured *C. trachomatis* IgG antibodies by microimmunofluorescence (MIF) assay, but we used a species-specific (peptide-based) immunoassay (EIA) as predictor of tubal pathology. These new EIA tests are well standardized, less laborious, less expensive, and easier to perform than the MIF test[16]. The disadvantage of using CAT is that there is yet no uniformity in assays and cut-off levels. In addition, *Chlamydia* antibody testing cannot predict tubal factor subfertility due to other causes (previous surgery or endometriosis), and the laboratory procedure has its limitations (e.g., cross-reactivity with other *Chlamydia* species such as *C. pneumoniae*). The greatest advantage of *Chlamydia* antibody testing is its simplicity, limited inconvenience, and no complications.

We demonstrated that *Chlamydia* antibody testing and HSG have comparable limited value in predicting pregnancy rates. No significant differences were found in cumulative pregnancy rate between CtsIgG-negative and CtsIgG-positive results. Idahl *et al*.[8] also showed that there were no differences between *C. trachomatis* IgG– negative or –positive women concerning pregnancy rates. There was also no difference in cumulative pregnancy rates between normal and abnormal findings at HSG. This low prognostic value of HSG was also found in other studies[9–12]. Only one recent prospective study showed results on the diagnostic and prognostic role of *Chlamydia* serology compared with HSG and laparoscopy and supported our findings[13]. That study did not record the definition of tubal pathology found at HSG or laparoscopy.

In conclusion, we assessed two diagnostic methods, CtsIgG and HSG, as a screening test for the likelihood of tubal damage or occurrence of pregnancy in subfertile women before laparoscopy. The diagnostic accuracy of *C. trachomatis*– specific IgG antibody testing is comparable with HSG, but both show poor performance. The prognostic value of occurrence of pregnancy of both tests is also poor. *Chlamydia* antibody testing as screening test to estimate the risk of tubal pathology before laparoscopy is preferable to HSG owing to its simplicity and limited inconvenience.

References

1. Hull MG, Glazener CM, Kelly NJ, Conway DI, Foster PA, Hinton RA, et al. Population study of causes, treatment, and outcome of infertility. Br Med J (Clin Res Ed) 1985;291:1693–7.
2. Rowe PJ, Comhaire FH, Hargreave TB and Mellows HJ. WHO manual for the standardized investigation and the diagnosis of the infertile couple. Cambridge: Cambridge University Press, 1993:4.
3. Dabekausen YAJM, Evers JLH, Land JA, Stals FS. *Chlamydia trachomatis* antibody testing is more accurate than hysterosalpingography in predicting tubal factor infertility. Fertil Steril 1994;61:833–7.
4. Meikle SF, Zhang X, Marine WM, Calonge BN, Hamman RF, Betz G. *Chlamydia trachomatis* antibody titers and hysterosalpingography in predicting tubal disease in infertility patients. Fertil Steril 1994; 62:305–12.
5. Mol BWJ, Dijkman B, Wertheim P, Lijmer J, van der Veen F, Bossuyt PMM. The accuracy of serum chlamydial antibodies in the diagnosis of tubal pathology: a meta-analysis. Fertil Steril 1997;67: 1031–7.
6. Thomas K, Couglin L, Mannion PT, Haddad NG. The value of *Chlamydia trachomatis* antibody testing as part of routine infertility investigation. Hum Reprod 2000;15:1079–82.
7. Veenemans LM, Linden PJ. The value of *Chlamydia trachomatis* antibody testing in predicting tubal factor infertility. Hum Reprod 2002;17:695– 8.
8. Idahl A, Boman J, Kumlin U, Olofsson JI. Demonstration of *Chlamydia trachomatis* IgG antibodies in the male partner of the infertile couple is correlated with a reduced likelihood of achieving pregnancy. Hum Reprod 2004;19:1121– 6.
9. Maas JWM, Evers JLH, ter Riet G, Kessels AGH. Pregnancy rate following normal versus abnormal hysterosalpingography findings: a meta-analysis. Gynecol Obstet Invest 1997;43:79–83.
10. Mol BW, Swart P, Bossuyt PM, van der Veen F. Is hysterosalpingography an important tool in predicting fertility outcome? Fertil Steril 1997;67:663–9.
11. Mol BW, Collins JA, Burrows EA, van der Veen F, Bossuyt PM. Comparison of hysterosalpingography and laparoscopy in predicting fertility outcome. Hum Reprod 1999;14:1237– 42.
12. Karande VC, Pratt DE, Rabin DS, Gleicher N. The limited value of hysterosalpingography in assessing tubal status and fertility potential. Fertil Steril 1995;63:1167–71.
13. Keltz MD, Gera PS, Moustakis M. *Chlamydia* serology screening in infertility patients. Fertil Steril 2006;85:752– 4.
14. Perquin DAM, Dörr PJ, de Craen AJM, Helmerhorst FM. The routine use of hysterosalpingography prior to laparoscopy in the fertility work-up: a multicenter randomized controlled trial. Hum Reprod 2006;21:1227–31.
15. Land JA, Evers JLH, Goossens VJ. How to use *Chlamydia* antibody testing in subfertility patients. Hum Reprod 1998;1094–8.
16. Verkooyen RP, Peeters MF, Rijsoort-Vos JH, van de Meijden WI, Mouton JW. Sensitivity and

specificity of three new commercially available *Chlamydia trachomatis* tests. Int J STD AIDS 2002;13: 23–5.

Chapter

5

Difficulties in recruitment for a randomized controlled trial involving hysterosalpingography

Denise AM Perquin[1], Anton JM de Craen[2] and Frans M Helmerhorst[3]

[1] Department of Obstetrics and Gynecology, Medical Center Haaglanden, The Hague, The Netherlands,

[2] Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands

[3] Department of Gynecology, Division of Reproductive Medicine, Leiden University Medical Center, Leiden, The Netherlands

Repod Health 2006 Jun 13; 3 : 5

Abstract

Background: The usefulness of hysterosalpingography (HSG) as routine investigation in the fertility work-up prior to laparoscopy and dye had been assessed in a randomized controlled trial. Recruiting subjects to the study was more difficult than anticipated. The objective of this study was to explore possible reasons for non-participation in the trial.

Methods: All newly referred subfertile women admitted to the Reproductive Medicine Clinic of Leiden University Medical Center between 1 April 1997 and 31 December 1999, were eligible for the study. The reasons for non-participation were evaluated by scrutinizing the medical records.

Results: Out of 759 women, a total of 127 (17%) agreed to participate in the trial. The most important reason for non-participation was because of exclusion criteria (73%). Other reasons were inattentive clinicians (3%) and patient-associated reasons (24%). Patient refusal and indecisiveness to enroll in the study were the most common patient-associated reasons. The most frequently stated reason for trial refusal was reluctance to undergo laparoscopy and dye mainly due to issues related to anesthesia and scheduling of procedure.

Conclusion: Almost three-quarters of recruitment difficulties in this study were due to unavoidable reasons. To overcome the remaining avoidable reasons for non-participation, attention should be paid to appropriate instruction of the study protocol to the participating doctors and to provide adequate information, in layman's terms, to the patients. Reminding patients by notes or telephone calls for attending the clinic are helpful. It may be contingent upon tracing the reasons of clinicians and patients for non-participation to improve enrollment during a trial.

Background

Between April 1997 and April 2002 we performed a pragmatic[1] multicenter randomized controlled trial comparing two different diagnostic strategies in the routine fertility work-up[2]. The hysterosalpingography (HSG) group underwent HSG first. If HSG showed normal uterine cavity, patent tubes and no tubal pathology, laparoscopy and dye followed after six months. In case of suspected tubal pathology, laparoscopy was performed within one or two months after HSG. The laparoscopy group did not receive HSG but underwent laparoscopy and dye directly after randomisation.

The power of the trial was based on randomisation of 750 subfertile women. Recruitment of patients into the trial was more difficult than expected. We estimated the highest recruitment rate from Leiden University Medical Center (LUMC). However, about halfway through the trial, we had only recruited 177 women instead of an estimated 375. To understand this low recruitment rate, we initiated the current study to find strategies to avoid the major reasons for non-participation which could be implemented during the second half of the study or later in other studies in reproductive medicine. This study identified potential eligible participants visiting one of the three participating hospitals (Leiden University Medical Center) during the first half of the recruitment period.

Methods

All women in our study participated in a multicenter randomized controlled trial with or without the performance of HSG to assess the usefulness of hysterosalpingography as routine investigation in the fertility work-up prior to laparoscopy and dye. Recruitment strategy, description of subjects, and main results of the trial have been published elsewhere[2]. In short, the trial was performed in one university hospital (Leiden University Medical Center, Leiden) and two non-university teaching hospitals (Medical Center Haaglanden, The Hague and Groene Hart Hospital, Gouda). All newly referred subfertile women who visited one of the three hospitals between April 1997 and April 2002 were eligible for inclusion in the trial. Exclusion criteria were a) subfertility less than 1 year, b) woman older than 37 years at time of first visit, c) anovulation in spite of clomifene citrate or bromocriptin use, d) abnormal semen analysis according to WHO criteria[3], or e) testing of tubal patency performed in the past. Women were asked to participate in the trial by their treating gynecologist at the time that HSG would normally be planned and informed consent was obtained. If the women refused to participate in the trial, the reason for non-

participation was recorded. A computer-based 1:1 ratio randomisation procedure was used to allocate the women into two groups; the HSG group or the laparoscopy group. Informed consent was obtained from all women. The Institutional Review Boards of each of the three hospitals approved all stages of the trial.
Recruitment of subjects into the trial was lower than expected. We elucidated this current study to explore the determinants of non-participation during the first half of the recruitment period among all potentially suitable subjects of Leiden University Medical Center to find strategies to avoid the major reasons for non-participation. We reviewed the medical records from all newly referred subfertile women who visited the Reproductive Medicine Clinic of Leiden University Medical Center from 1 April 1997 to 31 December 1999. The medical record of each subfertile couple contained either a sticker indicating that the woman participated in the study or documented the reasons for non-participation.

Results

From 1 April 1997 to 31 December 1999, 759 newly referred subfertile women visited the Reproductive Medicine clinic of Leiden University Medical Center. A total of 127 women (17%) met the inclusion criteria and agreed to participate in the trial, the remaining 632 did not. The unavoidable reasons (467 women; 73%) and avoidable reasons (165 women; 27%) for non-participation are summarized in table 1. Almost three-quarters of the women did not participate due to exclusion criteria (73%), 3% due to inattentive doctors and the remaining 153 women (24%) due to patient-associated reasons. From these 153 women, 72 of them refused and 19 women were indecisive to enroll in the study. Fifty women never showed up for randomisation after the initial visit. Personal circumstances such as leaving the area and relationship problems were also reported (n = 7).
Table 2 shows reasons for trial refusal among the 72 women. The most frequent stated reason was reluctance to laparoscopy and dye (35 women; 49%). Twenty-seven women (37%) did not state a reason for non-participation.

Table 1. Unavoidable and avoidable reasons for non-participation

Unavoidable reasons	N = 467	%
Exclusion criteria	460	73
Androgenic factor	173	
Tubal testing performed in the past	114	
Pregnant before randomisation	91	
Women older than 37 years at first visit	55	
Anovulation	27	
Patient's reason	7	1
Personal circumstances	7	
Avoidable reasons	N = 165	%
Doctor's reason	19	3
Eligible, but not approached	19	
Patient's reason	146	23
Refused to participate	72[1]	
No show-up after initial visit	50	
Indecisiveness	19	
Language barrier	5	

[1]see Table 2

Table 2: Patient's reasons for trial refusal

Reasons	N = 72	%
Reluctance to laparoscopy and dye:	35	49
General anesthesia	20	
The timing of the laparoscopy is too soon	15	
Reluctance to hysterosalpingography:	3	4
Fear for pain	3	
Don't want to be involved in a research project	7	10
Reasons not documented	27	37

Discussion

In retrospect, it seems clear that we had too optimistic recruitment targets. Main unavoidable reasons for nonparticipation in the trial were not meeting the inclusion criteria and personal circumstances. In 19 of 165 avoidable reasons for non-participation, it appeared that doctors were inattentive to approach their eligible patients for the trial. More details of their negligence were not documented, except that in general these doctors appeared to be willingly participating in the trial. Attention should be paid to appropriately instruct participating doctors in order to increase the recruitment of eligible patients. We have no evidence that physicians' preferences influenced the outcome of the randomized trial[4]. However, discussing the clinical relevance of the question as well as practical issues in the period that the protocol of the trial was designed appeared to be essential in the prevention of barriers in clinical recruitment[5].

Apprehensiveness towards one of the diagnostic procedures in this trial (laparoscopy) was mentioned by the women as the most prominent and avoidable reason for non-participation in the trial. General anesthesia prior to laparoscopy appeared to be a main obstacle for enrolment in the trial. Providing more adequate information on the actual procedure and using layman's terms may improve the rate of participation in such a trial. Although well educated and employed persons were more likely to refuse randomisation because of preference[4], we think that providing more information focused on problems that may emerge from questionnaires disseminated among potential participants in the development of the trial, may optimize recruitment. Some patients did not wish to be involved in a research project. Once patients have made up their mind and once they have prepared a distinct preference, it is nearly impossible to persuade them for enrolment[6].

One shortcoming of our paper is that we studied the major reasons of non-participation of potentially eligible participants visiting only one hospital. Unavoidable reasons of non-recruitment accounted for three quarters of the non-participation. The exclusion factor might be higher in an academic center due to specific criteria for referrals. The referred subfertile couples could have been older, with severe androgenic pathology or proven tubal pathology needing specialized assisted reproductive treatments. Another objective of this study was to find strategies to avoid the major reasons for non-participation which could be implemented during the second half of the study or later in other studies focusing on reproductive medicine. We assume that the major avoidable reasons of non-participation (like trial refusal) would be equally divided among all participating hospitals.

Planning for recruitment should be an important issue in the preparation period when a trial is designed[7,8]. Attention can also be paid to logistic problems that patients

may encounter. To minimize the no-show, reminder notes and telephone calls may remind patients to attend the clinic. A member of the research team, who can provide the information on a low profile with a high level of communication skills and understanding, can support the investigators. This person can deal with practical problems, such as patient's concerns or language barriers. This may contribute to solving the problem of women being less likely to participate in clinical trials[9].

Conclusion

In conclusion, this study showed that almost three-quarters of our recruitment failures were due to unavoidable reasons. To overcome the remaining avoidable reasons for non-participation and to increase external validity of a trial, it may be contingent upon tracing reasons of clinicians and patients for non-participation as well as by anticipating practical problems that clinicians and patients may encounter during a trial. In the set up of the trial and during the recruitment, communication and information are the key words.

Competing interests

The author(s) declare that they have no competing interests.

Authors' contributions

DAMP participated in the design of the study, collected and analyzed the data and wrote the first draft. AJMC performed statistical analyses and took part in the further preparation of the paper. FMH initiated the study, participated in the design of the study, took part in the further preparation of the paper and finalized the manuscript.

Acknowledgements

The authors wish to thank Kim Urgel for her help with collecting the data.

References

1. Haynes B: Can it work? Does it work? Is it worth it? The testing of healthcare interventions is evolving. BMJ 1999, 319:652-653.
2. Perquin DAM, Dorr PJ, de Craen AJM, Helmerhorst FM: Routine use of hysterosalpingography prior to laparoscopy in the fertility workup: a multicenter randomized controlled trial. Hum Reprod 2006, 21:1227-1231.
3. World Health Organization: WHO laboratory manual for the examination of human semen and semen-cervical mucus interaction 4th edition. Cambridge, Cambridge University Press; 1999:60-61.
4. King M, Nazareth I, Lampe F, Bower P, Chandler M, Morou M, Sibbald B, Lai R: Impact of participant and physician intervention preferences on randomized trials: a systematic review. JAMA 2005, 293:1089-1099.
5. Ross S, Grant A, Counsell C, Gillespie W, Russell I, Prescott R: Barriers to participation in randomised controlled trials: a systematic review. J Clin Epidemiol 1999, 52:1143-1156.
6. Mills N, Donovan JL, Smith M, Jacoby A, Neal DE, Hamdy FC: Perceptions of equipoise are crucial to trial participation: a qualitative study of men in the ProtecT study. Control Clin Trials 2003, 24:272-282.
7. Hunninghake DB, Darby CA, Probstfield JL: Recruitment experience in clinical trials: literature summary and annotated bibliography. Control Clin Trials 1987, 8(Suppl 4):6-30.
8. Collins JF, Bingham SF, Weiss DG, Williford WO, Kuhn RM: Some adaptive strategies for inadequate sample acquisition in veterans administration cooperative clinical trials. Control Clin Trials 1980, 1:227-248.
9. Recruitment of women to clinical trials. Lancet 2002, 358:853.

Chapter

6

Perinatal outcome of singletons and twins after assisted conception: a systematic review of controlled studies

Frans M Helmerhorst[1], Denise A.M. Perquin[1],
Diane Donker[1], Marc J.N.C. Keirse[2]

[1] Department of Gynaecology, Division of Reproductive Medicine, Leiden University Medical Center, Leiden, The Netherlands

[2] Department of Obstetrics, Gynaecology and Reproductive Medicine, Flinders University and Flinders Medical Center, Adelaide, South Australia

BMJ 2004;328:261

Abstract

Objective: To compare the perinatal outcome of singleton and twin pregnancies between natural and assisted conceptions.

Design: Systematic review of controlled studies published 1985-2002.

Studies reviewed: 25 studies were included of which 17 had matched and 8 had non-matched controls.

Main outcome measures: Very preterm birth, preterm birth, very low birth weight, low birth weight, small for gestational age, caesarean section, admission to neonatal intensive care unit, and perinatal mortality.

Results: For singletons, studies with matched controls indicated a relative risk of 3.27 (95% confidence interval 2.03 to 5.28) for very preterm (< 32 weeks) and 2.04 (1.80 to 2.32) for preterm (< 37 weeks) birth in pregnancies after assisted conception. Relative risks were 3.00 (2.07 to 4.36) for very low birth weight (< 1500 g), 1.70 (1.50 to 1.92) for low birth weight (< 2500 g), 1.40 (1.15 to 1.71) for small for gestational age, 1.54 (1.44 to 1.66) for caesarean section, 1.27 (1.16 to 1.40) for admission to a neonatal intensive care unit, and 1.68 (1.11 to 2.55) for perinatal mortality. Results of the non-matched studies were similar. In matched studies of twin gestations, relative risks were 0.95 (0.78 to 1.15) for very preterm birth, 1.07 (1.02 to 1.13) for preterm birth, 0.89 (0.74 to 1.07) for very low birth weight, 1.03 (0.99 to 1.08) for low birth weight, 1.27 (0.97 to 1.65) for small for gestational age, 1.21 (1.11 to 1.32) for caesarean section, 1.05 (1.01 to 1.09) for admission to a neonatal intensive care unit, and 0.58 (0.44 to 0.77) for perinatal mortality. The non-matched studies mostly showed similar trends.

Conclusions: Singleton pregnancies from assisted reproduction have a significantly worse perinatal outcome than non-assisted singleton pregnancies, but this is less so for twin pregnancies. In twin pregnancies, perinatal mortality is about 40% lower after assisted compared with natural conception.

Introduction

Twenty five years of assisted reproductive technology have not freed it from being a focus of medical, social, and political debate. Throughout this, reproductive technology has stood its ground, predominantly by offering parenthood to people who might not otherwise achieve it. However, issues that followed in its wake, such as surrogacy and pre-implantation diagnosis, have kept the momentum going on what many see as a loaded issue. It may be impossible to forecast where this will lead, but it should be possible to assess objectively whether babies born after assisted conception fare better or worse than those born after natural conception.

This question seems to be answered already by the widespread belief that pregnancy outcome is substantially worse after assisted conception.[1–3] The difference, however, relates predominantly to the higher frequency of multiple pregnancies.[3] The first indication that assisted singleton pregnancies may also have poorer outcomes appeared in 1985,[2] but it was not clear how much related to assisted reproduction or to confounders, such as maternal age and parity. Several matched cohort studies have since confirmed these findings.[1,4–8] Some studies found an opposite trend,[9,10] while most reported differences that were compatible with chance. Moreover, for twin pregnancies the general consensus, with few exceptions,[11–13] seems to be that assisted twin pregnancies have outcomes that are either similar to or slightly better than those conceived naturally.[1,9,14–17]

We identified all published studies on birth outcomes after assisted conception that distinguished singleton from multiple pregnancies and that incorporated an appropriate control group from the same population. We examined whether there are genuine differences in outcome between assisted and natural conceptions and whether they apply to both singleton and twin pregnancies.

Methods

We searched Medline, Embase, LILACS, and POPLINE for 1985-2002 with the MESH words perinatal care, fertilization in vitro, and the keywords perinatal outcome, perinatal care, assisted reproduction, and IVF, without language limitation. This search was supplemented with the references of the articles, review articles, and theses.

We selected reports with categorical data on any of the following outcomes: gestational age and weight at birth, caesarean section, perinatal death, and admission to neonatal intensive care. Studies without a control group of natural conceptions or that did not distinguish singleton from multiple pregnancies were excluded. The

remaining studies were controlled and we subdivided them into matched and non-matched, depending on the nature of the control group.
All authors read the studies, and at least two authors extracted data separately. Disagreements were resolved in discussion, if necessary after we contacted the original authors. All outcomes, except caesarean section, were expressed per number of infants.
International definitions were followed for preterm (< 37 weeks), very preterm (< 32 weeks), low birth weight (< 2500 g), very low birth weight (< 1500 g), small for gestational age (birth weight < 10th centile for gestation), and perinatal mortality (stillbirths and deaths in first week ≥ 500 g per 1000 total births ≥ 500 g). We used Review Manager (Update Software, Oxford) to calculate relative risks and 95% confidence intervals.

Results

Nine tables (A-I) of detailed results and a list of excluded studies can be found at the end of this chapter. Included studies are listed in tables A and B. Seventeen (14 matched and three non-matched) dealt with singleton pregnancies and 17 (10 matched and seven non-matched) with twin pregnancies. The tables show country and years covered by the study, types of assisted conception, number of cases, and type of controls.
Table 1 summarises relative risks of the outcomes in singleton and twin pregnancies after assisted and natural conception. Analyses for preterm birth and perinatal mortality are presented in tables 2-4, with more details and results for other outcomes in tables C-I. After table I the excluded studies with reasons for exclusion are listed.

Preterm birth
Very preterm singletons (< 32 weeks) were reported in only three studies with a prevalence of 1.3-2.1% in assisted conceptions and 0.3-2.9% in natural conceptions, a relative risk of 3.27 (95% confidence interval 2.03 to 5.28) (table 2).[1,9,19] Mildly preterm singletons (32-36 weeks) accounted for 6.5-9.2% and 3.8-7.6%, respectively, (2.05, 1.71 to 2.47) (see table C).[1,9,19]
Preterm singletons (< 37 weeks) accounted for 5.8-15% and 1.4-10.5%, respectively (table 2). The relative risk in both the 12 matched[1,4–10,12,19,21,22] and two non-matched[23,25] studies showed a doubling of the risk of preterm birth after assisted conception.
Very preterm twins were reported in three matched studies[1,9,19] (detailed in table 3) and two non-matched[17,26] studies. After we excluded one study that reported live

Table 1 Summary of risk of various outcomes in singleton and twin pregnancies after assisted conception compared with those conceived naturally. Figures are relative risk (95% confidence intervals)

	Singleton births		Twin births	
Outcome	**Matched studies**	**Non-matched studies**	**Matched studies**	**Non-matched studies**
Gestational age (weeks):				
<32	3.27 (2.03 to 5.28)	Not tested	0.95 (0.78 to 1.15)	1.20 (0.82 to 1.78)
32-36	2.05 (1.71 to 2.47)	Not tested	1.07 (1.00 to 1.14)	0.88 (0.66 to 1.17)
<37	2.04 (1.80 to 2.32)	1.94 (1.31 to 2.88)	1.07 (1.02 to 1.13)	0.99 (0.80 to 1.23)
Birth weight (g):				
<1500	3.00 (2.07 to 4.36)	1.57 (0.21 to 11.7)	0.89 (0.74 to 1.07)	1.46 (1.01 to 2.11)
1500-2499	1.54 (1.30 to 1.82)	3.28 (2.04 to 5.27)	1.02 (0.97 to 1.08)	1.05 (0.96 to 1.15)
<2500	1.70 (1.50 to 1.92)	2.58 (1.80 to 3.68)	1.03 (0.99 to 1.08)	1.12 (1.06 to 1.19)
Small for gestational age	1.40 (1.15 to 1.71)	1.46 (0.98 to 2.15)	1.27 (0.97 to 1.65)	0.93 (0.73 to 1.18)
Caesarean section	1.54 (1.44 to 1.66)	2.33 (1.95 to 2.79)	1.21 (1.11 to 1.32)	1.17 (1.06 to 1.29)
Neonatal intensive care unit	1.27 (1.16 to 1.40)	1.38 (0.67 to 2.86)	1.05 (1.01 to 1.09)	1.26 (1.16 to 1.36)
Perinatal mortality	1.68 (1.11 to 2.55)	3.77 (1.15 to 12.4)	0.58 (0.44 to 0.77)	0.84 (0.53 to 1.32)

Table 2 Preterm birth in singleton pregnancies after assisted conception compared with matched controls (natural conception)

Study	No (%) assisted	No (%) natural	Relative risk (95% CI)
Very preterm (<32 weeks)			
Dhont et al[1]*	63/3048 (2.1)	8/3048 (0.3)	7.88 (3.78 to 16.4)*
Dhont et al[9]	4/311 (1.3)	18/622 (2.9)	0.44 (0.15 to 1.30)
Koivurova et al[19]†	3/153 (2.0)	3/287 (1.0)	1.88 (0.38 to 9.18)†
Total	70/3512 (2.0)	29/3957 (0.7)	3.27 (2.03 to 5.28)
Preterm (<37 weeks)			
Dhont et al[1]*	344/3048 (11.3)	125/3048 (4.1)	2.75 (2.26 to 3.36)*
Dhont et al[9]	26/311 (8.4)	65/622 (10.5)	0.80 (0.52 to 1.23)
Isaksson et al[10]	4/69 (5.8)	35/345 (10.1)	0.57 (0.21 to 1.56)
Koivurova et al[19]†	13/153 (8.5)	16/287 (5.6)	1.52 (0.75 to 3.08)†
Koudstaal et al[8]	46/307 (15.0)	18/307 (5.9)	2.56 (1.52 to 4.30)
Nuojua et al[21]	8/92 (8.7)	14/276 (5.1)	1.71 (0.74 to 3.96)
Petersen et al[22]	5/70 (7.1)	3/70 (4.3)	1.67 (0.41 to 6.71)
Reubinoff et al[7]	23/260 (8.8)	10/260 (3.8)	2.30 (1.12 to 4.74)
Tallo et al[12]	6/62 (9.7)	1/62 (1.6)	6.00 (0.74 to 48.4)
Tan et al[4]	69/494 (14.0)	78/978 (8.0)	1.75 (1.29 to 2.38)
Tanbo et al[5]	53/355 (14.9)	61/643 (9.5)	1.57 (1.12 to 2.22)
Verlaenen et al[6]	16/140 (11.4)	2/140 (1.4)	8.00 (1.87 to 34.2)
Total	613/5361 (11.4)	428/7038 (6.1)	2.04 (1.80 to 2.32)

*Based on additional data obtained from authors.
†Surviving infants only.

Table 3 Preterm birth in twin pregnancies after assisted conception compared with matched controls (natural conception)

Study	No (%) assisted	No (%) natural	Relative risk (95% CI)
Very preterm (<32 weeks)			
Dhont et al[1]*	173/2482 (7.0)	178/2482 (7.2)	0.97 (0.79 to 1.19)*
Dhont et al[9]	16/230 (7.0)	12/230 (5.2)	1.33 (0.65 to 2.76)
Koivurova et al[19]†	2/103 (1.9)	11/103 (10.7)	0.18 (0.04 to 0.80)†
Total	191/2815 (6.8)	201/2815 (7.1)	0.95 (0.78 to 1.15)
Preterm (<37 weeks)			
Dhont et al[1]*	1227/2482 (49.4)	1184/2482 (47.7)	1.04 (0.98 to 1.10)*
Dhont et al[9]	120/230 (52.2)	98/230 (42.6)	1.22 (1.01 to 1.49)
Isaksson et al[10]	14/40 (35.0)	82/200 (41.0)	0.85 (0.54 to 1.34)
Koivurova et al[19]†	45/103 (43.7)	45/103 (43.7)	1.00 (0.73 to 1.36)†
Koudstaal et al[13]	98/192 (51.0)	80/192 (41.7)	1.23 (0.99 to 1.52)
Moise et al[11]	24/40 (60.0)	16/80 (20.0)	3.00 (1.81 to 4.98)
Petersen et al[22]	6/32 (18.8)	12/32 (37.5)	0.50 (0.21 to 1.17)
Tallo et al[12]†	40/68 (58.8)	25/68 (36.8)	1.60 (1.11 to 2.32)†
Tan et al[4]	146/250 (58.4)	22/42 (52.4)	1.11 (0.82 to 1.52)
Total	1720/3437 (50.0)	1564/3429 (45.6)	1.07 (1.02 to 1.13)

*Based on additional data obtained from the authors.
†Surviving infants only.

Table 4 Perinatal mortality in singleton and twin pregnancies after assisted conception compared with natural conception

Study	No (‰) assisted	No (‰) natural	Relative risk (95% CI)
Singletons			
Matched singleton studies:			
Dhont et al[1]*	41/3048 (13.5)	18/3048 (5.9)	2.28 (1.31 to 3.96)*
Dhont et al[9]	2/311 (6.4)	10/622 (16.1)	0.40 (0.09 to 1.81)
Isaksson et al[10]	1/69 (14.5)	5/345 (14.5)	1.00 (0.12 to 8.43)
Koudstaal et al[8]	3/307 (9.8)	1/307 (3.3)	3.00 (0.31 to 28.7)
Nuojua et al[21]	1/92 (10.9)	2/276 (7.2)	1.50 (0.14 to 16.4)
Reubinoff et al[7]	2/260 (7.7)	1/260 (3.8)	2.00 (0.18 to 21.9)
Tanbo et al[5]	4/355 (11.3)	6/643 (9.3)	1.21 (0.34 to 4.25)
Verlaenen et al[6]	3/140 (21.4)	2/140 (14.3)	1.50 (0.25 to 8.84)
Total	57/4582 (12.4)	45/5641 (8.0)	1.68 (1.11 to 2.55)
Non-matched singleton studies:			
Olivennes et al[25]†	3/162 (18.5)	25/5096 (4.9)	3.77 (1.15 to 12.4)
Twins			
Matched twin studies:			
Dhont et al[1]*	61/2482 (24.6)	82/2482 (33.0)	0.74 (0.54 to 1.03)*
Dhont et al[9]	0/230	6/230 (26.1)	0.08 (0.00 to 1.36)
Fitzsimmons et al[16]	4/112 (35.7)	48/216 (222.2)	0.16 (0.06 to 0.43)
Isaksson et al[10]	0/40	6/200 (30.0)	0.38 (0.02 to 6.56)
Koudstaal et al[13]	3/192 (15.6)	1/192 (5.2)	3.00 (0.31 to 28.6)
Tallo et al[12]	4/72 (55.6)	4/72 (55.6)	1.00 (0.26 to 3.85)
Total	72/3128 (23.0)	147/3392 (43.3)	0.58 (0.44 to 0.77)
Non-matched twin studies:			
Agustsson et al[14]	2/138 (14.5)	16/906 (17.6)	0.82 (0.19 to 3.53)
Lambalk et al[28]	18/1158 (15.5)	16/884 (18.1)	0.86 (0.44 to 1.67)
Olivennes et al[17]	10/144 (69.4)	28/328 (85.4)	0.81 (0.41 to 1.63)
Total	30/1440 (20.8)	60/2118 (28.3)	0.84 (0.53 to 1.32)

*Early neonatal deaths in this paper are erroneously labelled as early fetal deaths, but they are included in perinatal deaths.
†Two cases lost to follow up.

infants only,[19] the frequency range was 7.0-10.5% in assisted conceptions and 4.9-10.7% in natural conceptions and was not statistically different (see table D). Mildly preterm twins accounted for 41.7-45.2% of cases and 33.0-40.5% of controls in the matched studies (1.07, 1.00 to 1.14).[1,9,19] Preterm twins differed widely in frequency from 18.8-60.0% and 20.0-52.4%, respectively. The relative risk was 1.07 (1.02 to 1.13) in the nine matched studies[1,4,9–13,19,22] (table 3) and 0.99 (0.80 to 1.23) in the two non-matched studies (see table D).[17,23]

Birth weight

Singletons weighing < 1500 g were reported for six matched studies[1,5,6,9,10,19] and one non-matched study.[25] Frequencies in the matched studies were 1.5-3.9% for assisted conceptions and 0.32-7% for natural conceptions with a relative risk of 3.00 (2.07 to 4.36) (see table E). Singletons weighing < 2500 g were more common among cases than among controls in both matched (n = 12)[1,4–10,12,19,21,22] and non-matched (n =

2)[23,25] studies. Percentages of low birth weight were 2.9-15.7% in cases, 0-11.5% in matched controls, and 3.6-4.8% in non-matched controls (see table E).
Twins < 1500 g accounted for 5.0-25.0% of cases and 3.8-10.4% of controls (omitting one study reporting live infants only).[19] The relative risk was 0.89 (0.74 to 1.07) for the five matched[1,9–11,19] and 1.46 (1.01 to 2.11) for the two non-matched studies (see table F).[17,27] Twins < 2500 g accounted for 37.5-70.6% and 50.0-98.6% of cases versus 38.1-58.8% and 52.5-94.5% of controls, with relative risks of 1.03 (0.99 to 1.08) and 1.12 (1.06 to 1.19), respectively, in the eight matched studies[1,4,9–13,19,22] and the four non-matched studies.[17,23,26,27]

Small for gestational age
The 12 studies that reported on infants who were small for gestational age applied various reference charts. The frequency in singleton cases and controls was 1.6-16.3% versus 1.6-13.1% with a relative risk of 1.40 and 1.46, respectively, for the six matched[4,6–8,12,20] and two non-matched studies (see table G).[23,25] The four matched[4,11–13] and three non-matched[17,23,26] twin studies showed no significant difference between assisted and natural conceptions.

Caesarean section
Rates of caesarean section were significantly higher after assisted than after natural conception (see table H). The effect was more marked for singleton than for twin pregnancies in both matched[1,4–8,10,11,13,18,20,21] and non-matched studies.[14,17,24–27]

NICU admissions
Admissions to neonatal intensive care were more common after assisted conception in both matched and non-matched studies, and the difference was larger for singletons[1,5–7,9,10,19,21,23] than for twins (see table I).[1,9–11,14,19,23,26,27]

Perinatal mortality
Perinatal mortality differed widely among studies (table 4). In singleton pregnancies it was significantly higher after assisted than after natural conception in both matched and non-matched studies. All of the difference in the matched studies was accounted for by the study of Dhont et al in 1999, which contributed 67% of the cases.[1] Without this study mortality was 10.4 per 1000 for both cases and controls.
Matched twin studies were also dominated by the same study, which contributed 78% of the cases,[1] and by another with an extraordinarily high mortality among controls.[16] However, most twin studies showed a lower mortality after assisted than after natural conception, with a relative risk of 0.58 (0.44 to 0.77) for matched and 0.84 (0.53 to 1.32) for non-matched studies (table 4).

Discussion

Bias and confounding

Though assisted conception has had many successes, it seems that resulting singleton pregnancies have a worse outcome compared with naturally conceived singleton pregnancies. We chose to concentrate on birth issues and ignore early pregnancy outcomes, which are prone to ascertainment bias because they are detected more readily after assisted conception. This does not imply that birth outcomes are free from bias. Women with assisted pregnancies differ from other women in many characteristics that influence outcome, including age, parity, and socioeconomic status,[1,2,12] while subfertility itself also contributes to the difference.[29] We therefore subdivided studies into those with matched and those with non-matched, population specific controls and placed greater emphasis on the former. These virtually all matched for prominent confounders, such as age and parity, but they varied widely in controlling for other known confounders, such as socioeconomic status, smoking, and pre-existing disease. Although none controlled for all factors that might be important, they are likely to estimate true differences between assisted and natural conceptions better than the population based studies.

Nevertheless, our study uncovered major limitations of the matched cohort approach to differences in perinatal outcome between assisted and natural conceptions. Our summary results are largely dominated by a matched cohort study from Flanders, which contributed 54% of the cases in the singleton studies and 68% in the twin studies.[1] Its authors used three different control groups of singletons to match for various combinations of characteristics.[1] This led to disparate comparison groups, with perinatal mortality, for example, being 5.2 per 1000 in controls matched for maternal age and infant sex and 12.1 per 1000 in those matched also for parity and gestational age. The validity of matching for gestational age is questionable because gestational age is clearly influenced by assisted conception and affects other outcomes, such as birth weight and mortality. We therefore included only the controls matched for maternal age, infant sex, and parity. In another study, controls, but not cases, included several twin to twin transfusions in babies referred for special care.[16] Similar degrees of arbitrariness may have applied to other matched cohort studies without being apparent from the data.

Risk factors

Despite these limitations it is clear that the rate of preterm birth in singleton pregnancies after assisted reproduction is twice that seen with natural conceptions. This means that assisted reproduction is as much as a predictor for preterm birth as history of preterm birth.[30] The effect was larger for very preterm than for mildly

preterm births and translated into higher rates of (very) low birth weight, admission to intensive care, and perinatal death. However, not all of these should be attributed to preterm birth as there were also 40% more infants who were small for gestational age after assisted conception. There is some evidence that factors which influence gestational age at birth also influence weight for gestation,[31] and assisted conception may belong to the factors that influence both fetal weight and length of gestation.
On the other hand, if small for gestation fetuses are detected, this may prompt intervention that leads to earlier birth thereby contributing to both preterm and low birthweight rates. Unfortunately, we could not distinguish preterm births due to obstetric intervention from spontaneous preterm births. Neither could we ascertain that all singleton pregnancies, especially after assisted conception, were singleton pregnancies from the start rather than what remained after resorption of additional gestational sacs.

Twins

While the results of the matched and non-matched singleton studies invariably supported each other, this was less so for twin pregnancies. Differences between assisted and natural conceptions were all much smaller than in singleton pregnancies, often with confidence intervals that included unity. This is not due to smaller numbers because the overall sample size for twin studies was 84% of that of the singleton studies and the confidence intervals were smaller than for singletons.
An added risk, such as assisted conception, may have a marked impact on a low risk singleton pregnancy, but only a small effect on the heavily weighted balance of twin pregnancy. Assisted twin pregnancies may actually start off with a relative advantage over singleton pregnancies. As these studies were conducted when 85% of cycles of in vitro fertilisation entailed transfer of several embryos,[3,32] most births must have originated from the transfer of more than one embryo. Development of two rather than one may reflect an implantation advantage that accounts for the smaller difference in outcome between assisted and natural conceptions in twins than in singletons. Chorionicity certainly plays a part too. Dichorionic pregnancies fare better than monochorionic pregnancies and the latter account for 5-7% of assisted compared with 30% of natural twin pregnancies.[33] This effect was not obvious, though, in the studies that controlled for zygosity.[11,13,28] Earlier detection of twins with adaptation of antenatal care has been named as another factor,[17] but it is unclear what adaptations would significantly advantage assisted over natural twin pregnancies. However, none of this seems to explain the lower perinatal death rate in assisted than in natural twin pregnancies, especially as the other outcomes provide no indication how this might be mediated.

Conclusions and recommendations

Whatever the explanation may be, singletons from assisted conception are significantly disadvantaged compared with other singletons, but this is substantially less so for twins. Women undergoing assisted reproduction should be informed of the increased risks in singleton pregnancies. With a twin pregnancy they may be relatively advantaged compared with other twin gestations, but this is poor consolation for the much greater risks of twin pregnancy overall. Virtually all perinatal and infant morbidity occurs more frequently in twins than in singletons.[3]

Twenty five years after the birth of the first baby conceived by in vitro fertilisation, our data draw attention to three challenges. Firstly, emphasis needs to shift, more than it has already,[332] from achieving pregnancy to achieving a successful outcome. Secondly, it may be timely to consider any multiple pregnancy after assisted conception as a failure of that technology to achieve what it ought to achieve. Thirdly, there is a need to narrow the gap in perinatal outcome between assisted and other singleton pregnancies. This may also enhance understanding of how gestational age, fetal growth, and birth weight interact with each other.

References

1. Dhont M, De Sutter P, Ruyssinck G, Martens G, Bekaert A. Perinatal outcome of pregnancies after assisted reproduction: a case-control study. *Am J Obstet Gynecol* 1999;181:688-95.
2. Australian In Vitro Fertilisation Collaborative Group. High incidence of preterm births and early losses in pregnancy after in vitro fertilisation. *BMJ* 1985;291:1160-3.
3. Keirse MJNC, Helmerhorst FM. The impact of assisted reproduction on perinatal health care. *Soz Präventiv Med* 1995;40:343-51.
4. Tan SL, Doyle P, Campbell S, Beral V, Rizk B, Brinsden P, et al. Obstetric outcome of in vitro fertilization pregnancies compared with normally conceived pregnancies. *Am J Obstet Gynecol* 1992;167:778-84.
5. Tanbo T, Dale PO, Lunde O, Moe N, Abyholm T. Obstetric outcome in singleton pregnancies after assisted reproduction. *Obstet Gynecol* 1995;86:188-92.

6 Verlaenen H, Cammu H, Derde MP, Amy JJ. Singleton pregnancy after in vitro fertilization: expectations and outcome. *Obstet Gynecol* 1995;86:906-10.

7 Reubinoff BE, Samueloff A, Ben Haim M, Friedler S, Schenker JG, Lewin A. Is the obstetric outcome of in vitro fertilized singleton gestations different from natural ones? A controlled study. *Fertil Steril* 1997;67:1077-83.

8. Koudstaal J, Braat DD, Bruinse HW, Naaktgeboren N, Vermeiden JP, Visser GH. Obstetric outcome of singleton pregnancies after IVF: a matched control study in four Dutch university hospitals. *Hum Reprod* 2000;15:1819-25.

9. Dhont M, De Neubourg F, Van der Elst J, De Sutter P. Perinatal outcome of pregnancies after assisted reproduction: a case-control study. *J Assist Reprod Genet* 1997;14:575-80.
10. Isaksson R, Gissler M, Tiitinen A. Obstetric outcome among women with unexplained infertility after IVF: a matched case-control study. *Hum Reprod* 2002;17:1755-61.
11 Moise J, Laor A, Armon Y, Gur I, Gale R. The outcome of twin pregnancies after IVF. *Hum Reprod* 1998;13:1702-5.
12. Tallo CP, Vohr B, Oh W, Rubin LP, Seifer DB, Haning RV Jr. Maternal and neonatal morbidity associated with in vitro fertilization. *J Pediatr* 1995;127:794-800.
13. Koudstaal J, Bruinse HW, Helmerhorst FM, Vermeiden JP, Willemsen WN, Visser GH. Obstetric outcome of twin pregnancies after in-vitro fertilization: a matched control study in four Dutch university hospitals. *Hum Reprod* 2000;15:935-40.
14. Agustsson T, Geirsson RT, Mires G. Obstetric outcome of natural and assisted conception twin pregnancies is similar. *Acta Obstet Gynecol Scand* 1997;76:45-9.
15. Minakami H, Sayama M, Honma Y, Matsubara S, Koike T, Sato I, et al. Lower risks of adverse outcome in twins conceived by artificial reproductive techniques compared with spontaneously conceived twins. *Hum Reprod* 1998;13:2005-8.
16. Fitzsimmons BP, Bebbington MW, Fluker MR. Perinatal and neonatal outcomes in multiple gestations: assisted reproduction versus spontaneous conception. *Am J Obstet Gynecol* 1998;179:1162-7.
17. Olivennes F, Kadhel P, Rufat P, Fanchin R, Fernandez H, Frydman R. Perinatal outcome of twin pregnancies obtained after in vitro fertilization: comparison with twin pregnancies obtained spontaneously or after ovarian stimulation. *Fertil Steril* 1996;66:105-9.
18. D'Souza SW, Rivlin E, Cadman J, Richards B, Buck P, Lieberman BA. Children conceived by in vitro fertilisation after fresh embryo transfer. *Arch Dis Child Fetal Neonatal Ed* 1997;76:F70-4.
19. Koivurova S, Hartikainen AL, Gissler M, Hemminki E, Sovio U, Jarvelin MR. Neonatal outcome and congenital malformations in children born after in-vitro fertilization. *Hum Reprod* 2002;17:1391-8.
20. Maman E, Lunenfeld E, Levy A, Vardi H, Potashnik G. Obstetric outcome of singleton pregnancies conceived by in vitro fertilization and ovulation induction compared with those conceived spontaneously. *Fertil Steril* 1998;70:240-5.
21. Nuojua-Huttunen S, Gissler M, Martikainen H, Tuomivaara L. Obstetric and perinatal outcome of pregnancies after intrauterine insemination. *Hum Reprod* 1999;14:2110-5.
22. Petersen K, Hornnes PJ, Ellingsen S, Jensen F, Brocks V, Starup J, et al. Perinatal outcome after in vitro fertilisation. *Acta Obstet Gynecol Scand* 1995;74:129-31.
23. Addor V, Santos-Eggimann B, Fawer CL, Paccaud F, Calame A. Impact of infertility treatments on the health of newborns. *Fertil Steril* 1998;69:210-5.
24. Frydman R, Belaisch-Allart J, Fries N, Hazout A, Glissant A, Testart J. An obstetric assessment of the first 100 births from the in vitro fertilization program at Clamart, France. *Am J Obstet Gynecol* 1986;154:550-5.
25. Olivennes F, Rufat P, Andre B, Pourade A, Quiros MC, Frydman R. The increased risk of complication observed in singleton pregnancies resulting from in-vitro fertilization (IVF) does not seem to be

related to the IVF method itself. *Hum Reprod* 1993;8:1297-300.

26. Bernasko J, Lynch L, Lapinski R, Berkowitz RL. Twin pregnancies conceived by assisted reproductive techniques: maternal and neonatal outcomes. *Obstet Gynecol* 1997;89:368-72.
27. Daniel Y, Ochshorn Y, Fait G, Geva E, Bar-Am A, Lessing JB. Analysis of 104 twin pregnancies conceived with assisted reproductive technologies and 193 spontaneously conceived twin pregnancies. *Fertil Steril* 2000;74:683-9.
28. Lambalk CB, van Hooff M. Natural versus induced twinning and pregnancy outcome: a Dutch nationwide survey of primiparous dizygotic twin deliveries. *Fertil Steril* 2001;75:731-6.
29. Pandian Z, Bhattacharya S, Templeton A. Review of unexplained infertility and obstetric outcome: a 10 year review. *Hum Reprod* 2001;16:2593-7.
30. Keirse MJNC, Rush RW, Anderson ABM, Turnbull AC. Risk of pre-term delivery in patients with previous pre-term delivery and/or abortion. *Br J Obstet Gynaecol* 1978;85:81-5.
31. Keirse MJNC. International variations in intrauterine growth. *Eur J Obstet Gynecol Reprod Biol* 2000;92:21-8.
32. Nygren KG, Anderson AN. Assisted reproductive technology in Europe 1998. Results generated from European registers by ESHRE. *Hum Reprod* 2001;16:384-91.
33. Chow JS, Benson CB, Racowsky C, Doubilet PM, Ginsburg E. Frequency of a monochorionic pair in multiple gestations: relationship to mode of conception. *J Ultrasound Med* 2001;20:757-60.

Extra nine tables (A-I) of detailed results and list of excluded studies.

Table A Controlled singleton studies included in the review

Study	Years	Country	No. of cases	Type of cases*	Type of controls, matched for
Matched studies**					
Dhont et al. '97[9]	1991-1995	Belgium	311	1,2	Maternal age, parity, date of delivery (2 controls per case)
Dhont et al. '99[1]††	1992-1997	Belgium	3,048	1,2	Maternal age, parity, fetal sex, date of delivery
D'Souza et al.[18]	1984-1991	Great Britain	150	1	Maternal age, fetal sex, social class (controls are term babies)
Isaksson et al.[10]	1993-1999	Finland	69	1,2	Maternal age, parity, year of delivery, mother's residence (5 controls per case)
Koivurova et al.[19]‡	1990-1995	Finland	153	1	Maternal age, parity, fetal sex, year of delivery, area of residence, social class
Koudstaal et al.[8]	before 1993	Netherlands	307	1	Maternal age, parity, date of delivery
Maman et al.[20]	1989-1994	Israel	169	1	Maternal age, parity, gestational age, (2.8 controls per case)
Nuojua et al.[21]	1991-1996	Finland	92	4	Maternal age, parity, year of delivery, place of residence, smoking, occupation (3 controls per case)
Petersen et al.[22]	1986-1990	Denmark	70	1	Maternal age, parity
Reubinoff et al.[7]	1983-1993	Israel	260	1	Maternal age, parity, ethnic origin, location and date of delivery
Tallo et al.[12]	1988-1993	United States	62	1	Maternal age, race, insurance type, fetal sex, date of delivery
Tan et al.[4]	1978-1987	Great Britain	494	1	Maternal age
Tanbo et al.[5]	1985-1993	Norway	355	1,3,4	Maternal age, parity
Verlaenen et al.[6]	1988-1994	Belgium	140	1	Maternal age, parity, weight, height
Non-matched studies					
Addor et al.[23]	1993-1994	Switzerland	113	1,4	All women living in the canton of Vaud who gave birth to a singleton
Frydman et al.[24]	1981-1984	France	79	1	All spontaneous singletons delivered at the same hospital in the same period
Olivennes et al. '93[25]	1987-1989	France	164	1	Natural singletons delivered at the same institution in the same period

* Types of assisted conception studied: 1 = in-vitro fertilisation (IVF), 2 = intra-cytoplasmatic sperm injection (ICSI), 3 = gamete intra-fallopian transfer (GIFT), 4 = intra-uterine insemination (IUI).

** One control per case unless specified otherwise.

† Nine additional cases could not be matched accurately and are excluded from the analyses.

‡ Only surviving infants were included.

Table B Controlled twin studies included in the review

Study	Years	Country	No. of cases	Type of cases*	Type of controls, matched for
Matched studies**					
Dhont et al. '97[9]	1991-1995	Belgium	230	1,2	Maternal age, parity, date of delivery
Dhont et al. '99[1]†	1992-1997	Belgium	2,482	1,2	Maternal age, parity, date of delivery
Fitzsimmons et al.[16]	1985-1996	Canada	112	1,3,6	Maternal age, parity (two controls per case)
Isaksson et al.[10]	1993-1999	Finland	40	1,2	Maternal age, parity, year of delivery, mother's residence
Koivurova et al.[19]‡	1990-1995	Finland	103	1	Maternal age, parity, fetal sex, year of delivery, area of residence, parity, social class
Koudstaal et al.[13]	before 1993	The Netherlands	192	1	Maternal age, parity, date of delivery, zygosity
Moise et al.[11]	1990-1995	Israel	40	1	Maternal age, parity, ethnic origin, only dizygotic twins (2 controls per case)
Petersen et al.[22]	1986-1990	Denmark	32	1	Maternal age, parity
Tallo et al.[12]	1988-1993	United States	72	1	Maternal age, race, insurance type, fetal sex, date of delivery, birth order
Tan et al.[4]	1978-1987	Great Britain	250	1	Maternal age
Non-matched studies					
Addor et al.[23]	1993-1994	Switzerland	52	1,4	All women with naturally conceived twins living in the canton of Vaud
Agustsson et al.[14]	1990-1993	Iceland, Scotland	138	1	All naturally conceived twin pregnancies in Iceland and the Tayside Region, Scotland
Bernasko et al.[26]	1990-1995	United States	210	1,3	All naturally conceived twin deliveries in the database of Mount Sinai Medical Center
Daniel et al.[27]	1996-1997	Isra'l	208	1,2	All naturally conceived twins at Tel Aviv Sourasky Medical Center
Frydman et al.[24]	1981-1984	France	22	1	All naturally conceived twins delivered at the same hospital in the same period
Lambalk et al.[28]	1994-1996	The Netherlands	1,158	1,5	Primiparous dizygotic (male-female) twin deliveries
Olivennes et al. '96[17]	1988-1993	France	144	1	Naturally conceived twin pregnancies managed and delivered in Clamart

* Types of assisted conception studied: 1 = in-vitro fertilisation (IVF), 2 = intra-cytoplasmatic sperm injection (ICSI), 3 = gamete intra-fallopian transfer (GIFT), 4 = intra-uterine insemination (IUI), 5 = ovulation induction, 6 = clomiphene.

** One control per case unless specified otherwise.

† This paper also contains a subanalysis on 1,148 male-female twin infants matched for maternal age and parity, which is not included in this review.

‡ Only surviving infants were included.

Table C Preterm birth in singleton pregnancies after assisted *versus* natural conception

Study	Assisted		Spontaneous		Relative risk (95% CI)	
	n/N	%	n/N	%		
Very preterm (<32 weeks) in matched studies						
Dhont et al. '99[1] *	63/3,048	2.1	8/3,048	0.3	7.88	(3.78 to 16.4)*
Dhont et al. '97[9]	4/311	1.3	18/622	2.9	0.44	(0.15 to 1.30)
Koivurova et al.[19] **	3/153	2.0	3/287	1.0	1.88	(0.38 to 9.18)
Total very preterm (matched studies)	*70/3,512*	*2.0*	*29/3,957*	*0.7*	*3.27*	*(2.03 to 5.28)*
Mildly preterm (32-36 weeks) in matched studies						
Dhont et al. '99[1] *	281/3,048	9.2	117/3,048	3.8	2.40	(1.95 to 2.96)*
Dhont et al. '97[9]	22/311	7.1	47/622	7.6	0.94	(0.57 to 1.52)
Koivurova et al.[19] **	10/153	6.5	13/287	4.5	1.44	(0.65 to 3.21)**
Total mildly preterm (matched studies)	*313/3,512*	*9.1*	*177/3,957*	*4.5*	*2.05*	*(1.71 to 2.47)*
Preterm (<37 weeks) in matched studies						
Dhont et al. '99[1] *	344/3,048	11.3	125/3,048	4.1	2.75	(2.26 to 3.36)*
Dhont et al. '97[9]	26/311	8.4	65/622	10.5	0.80	(0.52 to 1.23)
Isaksson et al.[10]	4/69	5.8	35/345	10.1	0.57	(0.21 to 1.56)
Koivurova et al.[19] **	13/153	8.5	16/287	5.6	1.52	(0.75 to 3.08)**
Koudstaal et al.[8]	46/307	15.0	18/307	5.9	2.56	(1.52 to 4.30)
Nuojua et al.[21]	8/92	8.7	14/276	5.1	1.71	(0.74 to 3.96)
Petersen et al.[22]	5/70	7.1	3/70	4.3	1.67	(0.41 to 6.71)
Reubinoff et al.[7]	23/260	8.8	10/260	3.8	2.30	(1.12 to 4.74)
Tallo et al.[12]	6/62	9.7	1/62	1.6	6.00	(0.74 to 48.4)
Tan et al.[4]	69/494	14.0	78/978	8.0	1.75	(1.29 to 2.38)
Tanbo et al.[5]	53/355	14.9	61/643	9.5	1.57	(1.12 to 2.22)
Verlaenen et al.[6]	16/140	11.4	2/140	1.4	8.00	(1.87 to 34.2)
Total preterm (matched studies)	*613/5,361*	*11.4*	*428/7,038*	*6.1*	*2.04*	*(1.80 to 2.32)*
Preterm birth in non-matched studies						
Addor et al.[23]	6/113	5.3	280/6,088	4.6	1.15	(0.53-2.54)
Olivennes et al. '93[25] †	18/162	11.1	224/5,096	4.4	2.53	(1.61-3.98) †
Total preterm (non-matched studies)	*24/275*	*8.7*	*504/11,184*	*4.5*	*1.94*	*(1.31-2.88)*

* Based on additional data obtained from the authors. ** Only surviving infants. † Preterm in this study relates to <36 weeks.

Table D Preterm birth in twin pregnancies after assisted *versus* natural conception

Study	Assisted			Spontaneous		
	n/N	%		n/N	%	
Very preterm (<32 weeks)						
Matched studies						
Dhont et al. '99[1]*	173/2,482	7.0		178/2,482	7.2	
Dhont et al. '97[9]	16/230	7.0		12/230	5.2	
Koivurova et al.[19] †	2/103	1.9		11/103	10.7	
Total matched	*191/2,815*	*6.8*		*201/2,815*	*7.1*	
Non-matched studies						
Bernasko et al.[26]	22/210	10.5		56/558	10.0	
Olivennes et al. '96[17] **	12/144	8.3		16/328	4.9	
Total non-matched	*34/354*	*9.6*		*72/886*	*8.1*	
Mildly preterm (32-36 weeks)						
Matched studies						
Dhont et al. '99[1] *	1,054/2,482	42.5		1,006/2,482	40.5	
Dhont et al. '97[9]	104/230	45.2		86/230	37.4	
Koivurova et al.[19] †	43/103	41.7		34/103	33.0	
Total matched	*1,201/2,815*	*42.7*		*1,126/2,815*	*40.0*	
Non-matched studies						
Olivennes et al. '96[17] **	44/144	30.6		114/328	34.8	
Preterm (<37 weeks)						
Matched studies						
Dhont et al. '99[1] *	1,227/2,482	49.4		1,184/2,482	47.7	
Dhont et al. '97[9]	120/230	52.2		98/230	42.6	
Isaksson et al.[10]	14/40	35.0		82/200	41.0	
Koivurova et al[19] †	45/103	43.7		45/103	43.7	
Koudstaal et al.[13]	98/192	51.0		80/192	41.7	
Moise et al.[11]	24/40	60.0		16/80	20.0	
Petersen et al.[22]	6/32	18.8		12/32	37.5	
Tallo et al.[12] †	40/68	58.8		25/68	36.8	
Tan et al.[4]	146/250	58.4		22/42	52.4	
Total matched	*1,722/3,437*	*50.0*		*1,566/3,429*	*45.6*	
Non-matched studies						
Addor et al.[23]	12/26	46.2		69/154	44.8	
Olivennes et al. '96[17] **	56/144	38.9		130/328	39.6	
Total non-matched	*68/170*	*40.0*		*199/482*	*41.3*	

* Based on additional data obtained from the authors.

** Preterm relates to <36 weeks, very preterm to <31 weeks, and mildly preterm to 32-35 weeks; two cases were lost to follow-up.

† Surviving infants only.

	Relative risk (95% CI)	
	0.97	(0.79 to 1.19)*
	1.33	(0.65 to 2.76)
	0.18	(0.04 to 0.80)†
	0.95	*(0.78 to 1.15)*
	1.04	(0.65 to 1.67)
	1.71	(0.83 to 3.52)**
	1.20	*(0.82 to 1.78)*
	1.05	(0.98 to 1.12)*
	1.21	(0.97 to 1.51)
	1.26	(0.88 to 1.81) †
	1.07	*(1.00 to 1.14)*
	0.88	(0.66 to 1.17)**
	1.04	(0.98 to 1.10)*
	1.22	(1.01 to 1.49)
	0.85	(0.54 to 1.34)
	1.00	(0.73 to 1.36) †
	1.23	(0.99 to 1.52)
	3.00	(1.81 to 4.98)
	0.50	(0.21 to 1.17)
	1.60	(1.11 to 2.32)
	1.11	(0.82 to 1.52)
	1.07	*(1.02 to 1.13)*
	1.03	(0.66 to 1.62)
	0.98	(0.77 to 1.25)**
	0.99	*(0.80 to 1.23)*

Table E Low and very low birth weight in singleton pregnancies after assisted *versus* natural conception

Study	Assisted			Spontaneous		
	n/N	%		n/N	%	
Birthweight < 1,500 g						
Matched studies						
Dhont et al. '99[1]	72/3,048	2.4		10/3,048	0.3	
Dhont et al. '97[9]	5/311	1.6		17/622	2.7	
Isaksson et al.[10]	1/69	1.5		4/345	1.2	
Koivurova et al.[19]	3/153	2.0		2/287	0.7	
Tanbo et al.[5]	14/355	3.9		15/643	2.3	
Verlaenen et al.[6]	5/140	3.6		1/140	0.7	
Total matched	*100/4,076*	*2.5*		*49/5,085*	*1.0*	
Non-matched studies						
Olivennes et al. '93[25]	1/162	0.6		20/5,096	0.4	
Total non-matched	*1/162*	*0.6*		*20/5,096*	*0.4*	
Birthweight 1,500 - 2,499 g						
Matched studies						
Dhont et al. '99[1]	247/3,048	8.1		142/3,048	3.8	
Dhont et al. '97[9]	19/311	7.7		53/622	11.3	
Isaksson et al.[10]	1/69	1.4		13/345	3.8	
Koivurova et al.[19]	6/153	3.9		7/287	2.4	
Tanbo et al.[5]	27/355	7.6		28/643	4.4	
Verlaenen et al.[6]	9/140	6.4		5/140	3.6	
Total matched	*309/4,076*	*7.6*		*248/5,085*	*4.9*	
Non-matched studies						
Olivennes et al. '93[25]	17/162	10.5		163/5,096	3.2	
Total non-matched	*17/162*	*10.5*		*163/5,096*	*3.2*	
Birthweight < 2,500 g						
Matched studies						
Dhont et al. '99[1]	319/3,048	10.5		162/3,048	4.2	
Dhont et al. '97[9]	24/311	7.7		70/622	11.3	
Isaksson et al.[10]	2/69	2.9		17/345	5.0	
Koivurova et al.[19]	9/153	5.9		9/287	3.1	
Koudstaal et al.[8]	42/307	13.7		21/307	6.8	
Nuojua et al.[21]	8/92	8.7		17/276	6.2	
Petersen et al.[22]	11/70	15.7		0/70	0.0	
Reubinoff et al.[7]	29/260	11.2		30/260	11.5	
Tallo et al.[12]	3/62	4.8		3/62	4.8	
Tan et al.[4]	69/494	14.0		68/978	6.9	
Tanbo et al.[5]	41/355	11.5		43/643	6.7	
Verlaenen et al.[6]	14/140	10.0		6/140	4.3	
Total matched	*571/5,361*	*10.7*		*446/7,038*	*6.4*	
Non-matched studies						
Addor et al.[23]	11/113	9.7		292/6,088	4.8	
Olivennes et al. '93[25]	18/162	11.1		183/5096	3.6	
Total non-matched	*29/275*	*10.5*		*475/11,184*	*4.2*	

	Relative risk (95% CI)	
	7.20	(3.72 to 13.9)
	0.59	(0.22 to 1.58)
	1.25	(0.14 to 11.0)
	2.81	(0.48 to 16.7)
	1.69	(0.83 to 3.46)
	5.00	(0.59 to 42.3)
	3.00	*(2.07 to 4.36)*
	1.57	(0.21 to 11.7)
	1.57	*(0.21 to 11.7)*
	1.74	(1.42 to 2.13)
	0.72	(0.43 to 1.19)
	0.38	(0.05 to 2.89)
	1.61	(0.55 to 4.70)
	1.75	(1.05 to 2.92)
	1.80	(0.62 to 5.24)
	1.54	*(1.30 to 1.82)*
	3.28	(2.04 to 5.27)
	3.28	*(2.04 to 5.27)*
	1.97	(1.64 to 2.36)
	0.69	(0.44 to 1.07)
	0.59	(0.14 to 2.49)
	1.88	(0.76 to 4.63)
	2.00	(1.21 to 3.30)
	1.41	(0.63 to 3.16)
	23.0	(1.38 to 382.9)
	0.97	(0.60 to 1.56)
	1.00	(0.21 to 4.76)
	2.01	(1.46 to 2.76)
	1.73	(1.15 to 2.60)
	2.33	(0.92 to 5.90)
	1.70	*(1.50 to 1.92)*
	2.03	(1.14 to 3.60)
	3.09	(1.96 to 4.89)
	2.58	*(1.80 to 3.68)*

Table F Low and very low birth weight in twin pregnancies after assisted *versus* natural conception

Study	Assisted			Spontaneous		
	n/N	%		n/N	%	
Birthweight < 1,500 g						
Matched studies						
Dhont et al. '99[1]	166/2,482	6.7		196/2,482	7.9	
Dhont et al. '97[9]	14/230	6.1		14/230	6.1	
Isaksson et al.[10]	2/40	5.0		16/200	7.5	
Koivurova et al.[19] †	1/103	1.0		5/103	4.9	
Moise et al.[11]	10/40	25.0		3/80	3.8	
Total matched	*193/2,895*	*6.7*		*234/3,095*	*7.6*	
Non-matched studies						
Daniel et al.[27]	26/208	12.5		21/242	8.7	
Olivennes et al. '96[17]	22/144	15.3		34/328	10.4	
Total non-matched	*48/352*	*13.6*		*55/570*	*9.6*	
Birthweight 1,500 - 2,499 g						
Matched studies						
Dhont et al. '99[1]	1,182/2,482	47.6		1,157/2,482	46.6	
Dhont et al. '97[9]	114/230	49.6		112/230	48.7	
Isaksson et al.[10]	16/40	40.0		78/200	38.5	
Koivurova et al.[19] †	46/103	44.7		42/103	40.8	
Moise et al.[11]	18/40	45.0		44/80	55.0	
Total matched	*1,376/2,895*	*47.5*		*1,433/3,095*	*46.3*	
Non-matched studies						
Daniel et al.[26]	105/208	50.5		106/242	43.8	
Olivennes et al. '96[17]	120/144	83.3		276/328	84.1	
Total non-matched	*225/352*	*63.9*		*382/570*	*67.0*	
Birthweight < 2,500 g						
Matched studies						
Dhont et al. '99[1]	1,348/2,482	54.3		1,353/2,482	54.5	
Dhont et al. '97[9]	128/230	55.6		126/230	54.8	
Isaksson et al.[10]	18/40	45.0		92/200	46.0	
Koivurova et al.[19] †	47/103	45.6		47/103	45.6	
Koudstaal et al.[13]	117/192	60.9		85/192	44.3	
Moise et al.[11]	28/40	70.0		47/80	58.8	
Petersen et al.[22]	12/32	37.5		15/32	46.9	
Tallo et al.[12] †	48/68	70.6		29/68	42.6	
Tan et al.[4]	132/250	52.8		16/42	38.1	
Total matched	*1,878/3,437*	*54.6*		*1,810/3,429*	*52.8*	
Non-matched studies						
Addor et al.[23]	26/52	50.0		166/308	53.9	
Bernasko et al.[26] *	148/206	71.8		324/548	59.1	
Daniel et al.[27]	131/208	63.0		127/242	52.5	
Olivennes et al. '96[17]	142/144	98.6		310/328	94.5	
Total non-matched	*447/610*	*73.3*		*927/1,426*	*65.0*	

* Infants of 2 mothers in the study group and 5 in the control group were excluded because of stillbirth of one or both twins.

† Surviving infants only.

	Relative risk (95% CI)	
	0.85	(0.69 to 1.03)
	1.00	(0.49 to 2.05)
	0.63	(0.15 to 2.61)
	0.20	(0.02 to 1.68)
	6.67	(1.94 to 22.88)
	0.89	*(0.74 to 1.07)*
	1.44	(0.84 to 2.48)
	1.47	(0.89 to 2.43)
	1.46	*(1.01 to 2.11)*
	1.02	(0.96 to 1.08)
	1.02	(0.85 to 1.23)
	1.03	(0.68 to 1.56)
	1.10	(0.80 to 1.50)
	0.82	(0.55 to 1.22)
	1.02	*(0.97 to 1.08)*
	1.15	(0.95 to 1.40)
	0.99	(0.91 to 1.08)
	1.05	*(0.96 to 1.15)*
	1.00	(0.95 to 1.05)
	1.02	(0.86 to 1.20)
	0.98	(0.67 to 1.42)
	1.00	(0.74 to 1.35)
	1.38	(1.13 to 1.67)
	1.19	(0.91 to 1.57)
	0.80	(0.45 to 1.43)
	1.66	(1.21 to 2.27)
	1.39	(0.93 to 2.07)
	1.03	*(0.99 to 1.08)*
	0.93	(0.69 to 1.24)
	1.22	(1.09 to 1.36)
	1.20	(1.02 to 1.41)
	1.04	(1.01 to 1.08)
	1.12	*(1.06 to 1.19)*

Table G Small for gestational infants from singleton and twin pregnancies after assisted conception *versus* natural conceptions*

Study*	Assisted			Spontaneous			Relative risk (95% CI)	
	n/N	%		n/N	%			
Singletons								
Matched studies								
Koudstaal et al.[8]	50/307	16.3		24/307	7.8		2.08	(1.31 to 3.30)
Maman et al.[20]	13/169	7.7		36/469	7.7		1.00	(0.54 to 1.84)
Reubinoff et al.[7]	33/260	12.7		34/260	13.1		0.97	(0.62 to 1.52)
Tallo et al.[12]	1/62	1.6		1/62	1.6		1.00	(0.06 to 15.63)
Tan et al.[4]	64/494	13.0		89/978	9.1		1.42	(1.05 to 1.93)
Verlaenen et al.[6]	15/140	10.7		7/140	5.0		2.14	(0.90 to 5.09)
Total matched	*176/1,432*	*12.3*		*191/2,216*	*8.6*		*1.40*	*(1.15 to 1.71)*
Non-matched studies								
Addor et al.[23]	6/113	5.3		371/6,088	6.1		0.87	(0.40 to 1.91)
Olivennes et al. '93[25]	18/162	11.1		301/5,096	5.9		1.88	(1.20 to 2.95)
Total non-matched	*24/275*	*8.7*		*672/11,184*	*6.0*		*1.46*	*(0.98 to 2.15)*
Twins								
Matched studies								
Koudstaal et al.[13]	51/192	26.6		49/192	25.5		1.04	(0.74 to 1.46)
Moise et al.[11]	12/40	30.0		15/80	18.8		1.60	(0.83 to 3.09)
Tallo et al.[12] †	13/68	19.1		7/68	10.3		1.86	(0.79 to 4.37)
Tan et al.[4]	58/250	23.2		6/42	14.3		1.62	(0.75 to 3.52)
Total matched	*134/550*	*24.4*		*77/382*	*20.2*		*1.27*	*(0.97 to 1.65)*
Non-matched studies								
Addor et al.[23]	16/52	30.8		92/308	29.9		1.03	(0.66 to 1.60)
Bernasko et al.[26] **	26/206	12.6		68/548	12.4		1.02	(0.67 to 1.55)
Olivennes et al. '96[17]	26/144	18.1		75/328	22.9		0.79	(0.53 to 1.18)
Total non-matched	*68/402*	*16.9*		*235/1,184*	*19.8*		*0.93*	*(0.73 to 1.18)*

* Different weight for gestational age curves were used as references in these studies. Koudstaal et al.[8,13] and Verlaenen et al.[6] used the reference curve of Kloosterman, Reubinoff et al.[7] and Tan et al.[4] the Gairdner and Pearson standard growth charts, Olivennes et al.[17, 25] the Leroy and Lefort curve, and Moise et al.[11] and Tallo et al.[12] the Colorado Intrauterine Growth Charts. Maman et al.[20], Addor et al.[23] and Bernasko et al.[26] did not report what standard charts were used.

** Infants of 2 mothers in the study group and 5 in the control group were excluded from the report because of stillbirth of one or both twins.

† Surviving infants only.

Table H Caesarean section rates in singleton and twin pregnancies after assisted *versus* natural conceptions

Study	Assisted			Spontaneous			Relative risk (95% CI)	
	n/N	%		n/N	%			
Singletons								
Matched studies								
Dhont et al. '99[1]	647/3,048	21.2		496/3,048	16.3		1.30	(1.17 to 1.45)
D'Souza[18]	40/150	26.7		19/150	12.7		2.11	(1.28 to 3.46)
Isaksson et al.[10]	17/69	24.6		70/345	20.3		1.21	(0.76 to 1.93)
Koudstaal et al.[8]	51/307	16.6		40/307	13.0		1.27	(0.87 to 1.87)
Maman et al.[20]	80/169	47.3		93/469	19.8		2.39	(1.87 to 3.04)
Nuojua et al.[21]	23/92	25.0		69/276	25.0		1.00	(0.66 to 1.50)
Reubinoff et al.[7]	109/260	41.9		40/260	15.4		2.73	(1.98 to 3.75)
Tan et al.[4]	232/494	47.0		235/978	24.0		1.95	(1.69 to 2.26)
Tanbo et al.[5]	100/355	28.2		125/643	19.4		1.45	(1.15 to 1.82)
Verlaenen et al.[6]	16/140	11.4		10/140	7.1		1.60	(0.75 to 3.40)
Total matched	*1,315/5,084*	*25.9*		*1,197/6,616*	*18.1*		1.54	(1.44 to 1.66)
Non-matched studies								
Frydman et al.[24]	37/79	46.8		597/3,841	15.5		3.01	(2.36 to 3.85)
Olivennes et al. '93[25]	47/162	29.0		749/5,096	14.7		1.97	(1.54 to 2.53)
Total non - matched	*84/241*	*34.9*		*1,346/8,937*	*15.1*		*2.33*	*(1.95 to 2.79)*
Twins								
Matched studies								
Dhont et al. '99[1]	559/1,241	45.0		457/1,241	36.8		1.22	(1.11 to 1.35)
Isaksson et al.[10]	12/20	60.0		68/100	68.0		0.88	(0.60 to 1.29)
Koudstaal et al.[13]	38/96	39.6		29/96	30.2		1.31	(0.89 to 1.94)
Moise et al.[11]	13/20	65.0		21/40	52.5		1.24	(0.80 to 1.92)
Tan et al.[4]	80/125	64.0		13/21	61.9		1.03	(0.72 to 1.48)
Total matched	*702/1,502*	*46.7*		*588/1,498*	*39.3*		*1.21*	*(1.11 to 1.32)*
Non-matched studies								
Agustsson et al.[14]	45/69	65.2		259/453	57.2		1.14	(0.94 to 1.38)
Bernasko et al.[26]	53/105	50.5		142/279	50.9		0.99	(0.79 to 1.24)
Daniel et al.[27]	45/104	43.3		41/121	33.9		1.28	(0.92 to 1.78)
Frydman et al.[24]	8/11	72.7		26/58	44.8		1.62	(1.02 to 2.57)
Lambalk et al.[28]	181/480	37.7		192/613	31.3		1.20	(1.02 to 1.42)
Olivennes et al. '96[17]	39/72	54.2		71/164	43.3		1.25	(0.95 to 1.65)
Total non- matched	*371/841*	*44.1*		*731/1,688*	*43.3*		*1.17*	*(1.06 to 1.29)*

Table I Admissions to neonatal intensive care units after singleton and twin pregnancies from assisted *versus* natural conceptions

Study	Assisted			Spontaneous			Relative risk (95% CI)	
	n/N	%		n/N	%			
Singletons								
Matched studies								
Dhont et al. '99[1]	638/3,048	20.9		506/3,048	16.6		1.26	(1.13 to 1.4)
Dhont et al. '97[9]	30/311	9.6		61/622	9.8		0.98	(0.65 to 1.49)
Isaksson et al.[10]	1/69	1.5		0/345	0.0		14.8	(0.61 to 360)
Koivurova et al.[19] †	20/153	13.1		30/287	10.5		1.25	(0.74 to 2.13)
Nuojua et al.[21]	2/92	2.2		18/276	6.5		0.33	(0.08 to 1.41)
Reubinoff et al.[7]	9/260	3.5		5/260	1.9		1.80	(0.61 to 5.30)
Tanbo et al.[5]	42/355	11.8		40/643	6.2		1.90	(1.26 to 2.87)
Verlaenen et al.[6]	23/140	16.4		14/140	10.0		1.64	(0.88 to 3.06)
Total matched	*765/4428*	*17.3*		*674/5,621*	*12.0*		*1.27*	*(1.16 to 1.40)*
Non-matched studies								
Addor et al.[23]	7/113	6.2		273/6,088	4.5		1.38	(0.67 to 2.86)
Total non-matched	*7/113*	*6.2*		*273/6,088*	*4.5*		*1.38*	*(0.67 to 2.86)*
Twins								
Matched studies								
Dhont et al. '99[1]	1,727/2,482	69.6		1,682/2,482	67.8		1.03	(0.99 to 1.07)
Dhont et al. '97[9]	98/230	42.6		60/230	26.1		1.63	(1.25 to 2.13)
Isaksson et al.[10]	4/40	10.0		16/200	7.5		1.25	(0.44 to 3.54)
Koivurova et al.[19] †	39/103	37.9		46/103	44.7		0.85	(0.61 to 1.18)
Moise et al.[11]	16/40	40.0		23/80	28.8		1.39	(0.83 to 2.32)
Total matched	*1,884/2,895*	*65.1*		*1,827/3,095*	*59.0*		*1.05*	*(1.01 to 1.09)*
Non-matched studies								
Addor et al.[23]	14/52	26.9		38/308	12.3		2.18	(1.27 to 3.74)
Agustsson et al.[14]	126/138	91.3		736/906	81.2		1.12	(1.06 to 1.19)
Bernasko et al.[26] *	96/206	46.6		200/548	36.5		1.28	(1.06 to 1.53)
Daniel et al.[27]	93/208	44.7		75/242	30.9		1.44	(1.13 to 1.84)
Total non-matched	*329/604*	*54.5*		*1,049/2,004*	*52.3*		*1.26*	*(1.16 to 1.36)*

* Infants of 2 mothers in study group and 5 in the control group excluded because of stillbirth of one or both twins.

† Denominators include only surviving infants.

References to studies examined, but excluded from the analyses (with reasons for their exclusion)

Australian in vitro fertilisation collaborative group. High incidence of preterm births and early losses in pregnancy after in vitro fertilisation. *BMJ* 1985; 291:1160-63. (overlap with Saunders *et al.* and no control group)

Australian In-Vitro Fertilization Collaborative Group. In-vitro fertilization pregnancies in Australia and New Zealand, 1979-1985. *Med J Aust* 1988;148:429-36. (overlap with Saunders *et al.* and no control group)

MRC Working Party on Children Conceived by In Vitro Fertilisation. Births in Great Britain resulting from assisted conception, 1978-87. *BMJ* 1990;300:1229-33. (no control group)

FIVNAT (French In Vitro National). Pregnancies and births resulting from in vitro fertilization: French national registry, analysis of data 1986 to 1990. *Fertil Steril* 1995;64:746-56. (no control group)

Andrews MC, Muasher SJ, Levy DL, Jones HW, Jr., Garcia JE, Rosenwaks Z et al. An analysis of the obstetric outcome of 125 consecutive pregnancies conceived in vitro and resulting in 100 deliveries. *Am J Obstet Gynecol* 1986;154:848-54. (no separate data for singleton and twin pregnancies)

Beral V, Doyle P, Tan SL, Mason BA, Campbell S. Outcome of pregnancies resulting from assisted conception. *Br Med Bull* 1990;46:753-68. (no control group)

Bergh T, Ericson A, Hillensjo T, Nygren KG, Wennerholm UB. Deliveries and children born after in-vitro fertilisation in Sweden 1982-95: a retrospective cohort study. *Lancet* 1999;354:1579-785. (no control group)

Brandes JM, Scher A, Itzkovits J, Thaler I, Sarid M, Gershoni-Baruch R. Growth and development of children conceived by in vitro fertilization. *Pediatrics* 1992;90:424-9. (no specific obstetric data)

Callahan TL, Hall JE, Ettner SL, Christiansen CL, Greene MF, Crowley WF, Jr. The economic impact of multiple-gestation pregnancies and the contribution of assisted-reproduction techniques to their incidence. *N Engl J Med* 1994;331:244-9. (no specific obstetric data)

Cohen J, Mayaux MJ, Guihard-Moscato ML. Pregnancy outcomes after in vitro fertilization. A collaborative study on 2342 pregnancies. Ann N Y Acad Sci 1988;541:1-6. (no control group)

Doyle P, Beral V, Maconochie N. Preterm delivery, low birthweight and small-for-gestational-age in

liveborn singleton babies resulting from in-vitro fertilization. Hum Reprod 1992;7:425-8. (no numerical data for the control group reported to be “the general population of England and Wales”)

Friedler S, Mashiach S, Laufer N. Births in Israel resulting from in-vitro fertilization/embryo transfer, 1982-1989: National Registry of the Israeli Association for Fertility Research. *Hum Reprod* 1992;7:1159-63. (no control group)

Ghazi HA, Spielberger C, Kallen B. Delivery outcome after infertility--a registry study. *Fertil Steril* 1991;55:726-32. (no separate data for singleton and twin pregnancies and no control group)

Gissler M, Malin SM, Hemminki E. In-vitro fertilization pregnancies and perinatal health in Finland 1991-1993. *Hum Reprod* 1995;10:1856-61. (no control group)

Hack M, Brish M, Serr DM, Insler V, Salomy M, Lunenfeld B. Outcome of pregnancy after induced ovulation. Follow-up of pregnancies and children born after clomiphene therapy. *JAMA* 1972;220:1329-33. (no assisted procreation and no control group)

Jonas HA, Lumley J. Triplets and quadruplets born in Victoria between 1982 and 1990. The impact of IVF and GIFT on rising birthrates. *Med J Aust* 1993;158:659-63. (no separate data for singletons and twins)

Lancaster PA. Obstetric outcome. *Clin Obstet Gynaecol* 1985;12:847-64. (no specific obstetric data and overlap with Australian In-Vitro Fertilization Collaborative Group, 1985)

Levene MI, Wild J, Steer P. Higher multiple births and the modern management of infertility in Britain. The British Association of Perinatal Medicine. *Br J Obstet Gynaecol* 1992;99:607-13. (no separate data for singletons and twins and no control group)

Li TC, MacLeod I, Singhal V, Duncan SL. The obstetric and neonatal outcome of pregnancy in women with a previous history of infertility: a prospective study. *Br J Obstet Gynaecol* 1991;98:1087-92. (data not related to the aim of our study)

McFaul PB, Patel N, Mills J. An audit of the obstetric outcome of 148 consecutive pregnancies from assisted conception: implications for neonatal services. *Br J Obstet Gynaecol* 1993;100:820-5. (no specific obstetric data and no control group)

Minakami H, Sayama M, Honma Y, Matsubara S, Koike T, Sato I et al. Lower risks of adverse outcome in twins conceived by artificial reproductive techniques compared with spontaneously conceived twins. *Hum Reprod* 1998;13:2005-8. (no specific obstetric data)

Rizk B, Doyle P, Tan SL, Rainsbury P, Betts J, Brinsden P et al. Perinatal outcome and congenital malformations in in-vitro fertilization babies from the Bourn-Hallam group. *Hum Reprod* 1991;6:1259-64. (part of MRC Working Party 1990 and no control group)

Rufat P, Olivennes F, de Mouzon J, Dehan M, Frydman R. Task force report on the outcome of pregnancies and children conceived by in vitro fertilization (France: 1987 to 1989). *Fertil Steril* 1994;61:324-30. (overlap with FIVNAT and no control group)

Sassoon DA, Castro LC, Davis JL, Hobel CJ. Perinatal outcome in triplet versus twin gestations. *Obstet Gynecol* 1990;75:817-20. (no control group)

Saunders DM,.Lancaster P. The wider perinatal significance of the Australian in vitro fertilization data collection program. *Am J Perinatol* 1989;6:252-7. (no control group)

Seoud MA, Toner JP, Kruithoff C, Muasher SJ. Outcome of twin, triplet, and quadruplet in vitro fertilization pregnancies: the Norfolk experience. *Fertil Steril* 1992;57:825-34. (no specific obstetric data

Schieve LA.,Meikle SF, Ferre C, Peterson HB, Jeng G, Wilcox LS. Low and very low birth weight in infants conceived with use of assisted reproductive technology. *New Engl J Med* 2002;346:731-7. (no numerical outcome data for control group; based on birth notifications)

Spellacy WN, Handler A, Ferre CD. A case-control study of 1253 twin pregnancies from a 1982-1987 perinatal data base. *Obstet Gynecol* 1990;75:168-71. (no assisted procreation and no control group)

Steptoe PC, Edwards RG, Walters DE. Observations on 767 clinical pregnancies and 500 births after human in-vitro fertilization. *Hum Reprod* 1986;1:89-94. (no control group)

Tan SL, Maconochie N, Doyle P, Campbell S, Balen A, Bekir J *et al*. Cumulative conception and live-birth rates after in vitro fertilization with and without the use of long, short, and ultrashort regimens of the gonadotropin-releasing hormone agonist buserelin. *Am J Obstet Gynecol* 1994;171:513-20. (no specific obstetric data and no control group)

Tanbo T,.Abyholm T. Obstetric and perinatal outcome in pregnancies after assisted reproduction. *Curr Opin Obstet Gynecol* 1996;8:193-8. (no control group)

Tuck SM, Yudkin PL, Turnbull AC. Pregnancy outcome in elderly primigravidae with and without a history of infertility. *Br J Obstet Gynaecol* 1988;95:230-7. (no assisted procreation)

Venn A,.Lumley J. Clomiphene citrate and pregnancy outcome. *Aust N Z J Obstet Gynaecol* 1994;34:56-66. (no assisted procreation)

Wang JX, Clark AM, Kirby CA, Philipson G, Petrucco O, Anderson G *et al*. The obstetric outcome of singleton pregnancies following in-vitro fertilization/gamete intra-fallopian transfer. *Hum Reprod* 1994;9:141-6. (no specific obstetric data)

Wennerholm UB, Janson PO, Wennergren M, Kjellmer I. Pregnancy complications and short-term follow-up of infants born after in vitro fertilization and embryo transfer (IVF/ET). *Acta Obstet Gynecol Scand* 1991;70:565-573. (no control group)

Westergaard T, Wohlfahrt J, Aaby P, Melbye M. Population based study of rates of multiple pregnancies in Denmark, 1980-94. *BMJ* 1997;314:775-9. (no assisted procreation)

Williams MA, Goldman MB, Mittendorf R, Monson RR. Subfertility and the risk of low birth weight. *Fertil Steril* 1991;56:668-71. (no separate data on singletons and twins and no control group)

Wisanto A, Magnus M, Bonduelle M, Liu J, Camus M, Tournaye H *et al*. Obstetric outcome of 424 pregnancies after intracytoplasmic sperm injection. *Hum Reprod* 1995;10:2713-8. (no specific obstetric data and no control group)

Yeh J, Leipzig S, Friedman EA, Seibel MM. Results of in vitro fertilization pregnancies: experience at Boston's Beth Israel Hospital. *Int J Fertil* 1990;35:116-9. (overlap with Friedler *et al.* and no control group)

Chapter 7

General discussion

7. General introduction

The first aim of this thesis was to assess the value of hysterosalpingography (HSG) prior to laparoscopy and dye in a routine clinical setting. The results of the randomized controlled trial comparing fertility work-ups with or without HSG will be discussed in the first part of this chapter. The second aim was to clarify whether adverse perinatal outcome is related to assisted conception or other maternal risk factors and whether this is for both singleton and twin pregnancies. The results of the second study on this subject will be discussed in the last part of this chapter.

7.1 Evaluation of diagnostic strategies in reproductive medicine: a study on hysterosalpingography

7.1.1 Study objective and methodology

Routine clinical investigation of subfertile couple consists, apart from ovulation detection and sperm analysis, of tubal patency detection. The first objective of this thesis was to investigate whether hysterosalpingography (HSG) prior to laparoscopy and dye should be maintained as a diagnostic procedure in the routine fertility work-up.

Assessment of tubal function is a fundamental part of the diagnostic fertility work-up. HSG has become the first-line approach to assess the patency of the fallopian tubes in routine fertility work-up due to the non-invasive, outpatient procedure and low cost[1,2] although laparoscopy and dye is the reference standard according to WHO criteria[3]. A reason for performing HSG in stead of or prior to laparoscopy and dye cannot be found in the test characteristics of HSG but is based more on tradition and personal preference[1]. Many authors have questioned the validity of HSG in the diagnosis of tubal function due to its moderate sensitivity and specificity compared to laparoscopy and dye[4]. Although flushing of the tubes with oil-soluble contrast media during HSG has been reported to shorten the period to become pregnant[5], many women experience HSG as a painful procedure. The availability of other investigations for assessing tubal function like Chlamydia antibody testing or other new methods (hysterosalpingo-contrast-sonography, falloscopy, transvaginal laparoscopy) have provided new modalities to evaluate the tubal function. There are many arguments why HSG should or should not be part of the standard diagnostic fertility work-up based on its diagnostic value, pregnancy rates, pain, costs and alternatives.

Because the usefulness of HSG was uncertain, a pragmatic multicenter randomized controlled trial comparing fertility work-up with or without HSG to assess the value

of HSG prior to laparoscopy and dye in a routine clinical setting was carried out. A pragmatic effectiveness trial was designed to evaluate whether two commonly used diagnostic strategies (HSG and laparoscopy & dye) have a different impact on cumulative pregnancy rate of subfertile women (chapter 2). A similar cumulative pregnancy rate at 18 months in the intervention (HSG) group (49.1%) and the control (non-HSG) group (50.3%) was found. We concluded that routine use of HSG at an early stage in the fertility work-up prior to laparoscopy and dye does not influence cumulative pregnancy rate compared to routine use of laparoscopy without HSG.

The design of the trial and thus our conclusion was criticized by Coppus *et al.*[6]. They argued that clinical outcomes are not affected by the evaluated test but by the intervention that follows a normal or abnormal test result. As such, the randomized comparisons of the use of diagnostic tests or strategies could only show clinical benefits if an abnormal test result is followed by a standardized and explicit treatment protocol, according the authors of the letter[6]. However, the suggestion to change the design of the trial by only randomizing patients with discordant test results would have resulted in an *efficacy* trial and would have evaluated the effect of the test under ideal circumstances. In such a trial the question 'can it work?' is answered[7]. Our trial, however, was an *effectiveness* trial designed to compare two common clinical diagnostic strategies in the setting of daily clinical practice. Hence, in our trial the question 'does it work in practice?' was answered[7]. In that setting, a worthwhile test plus intervention must have an effect that is robust enough to overcome differences in patient choice, local procedures and health care access.

7.1.2 Pain and patient evaluation

Two studies have shown that women may experience HSG as a painful procedure[8,9]. Laparoscopy is invasive and generally needs general anaesthesia. In an additional analysis of subfertile women nested in the above mentioned randomized controlled trial, pain and patient satisfaction of women undergoing an HSG or laparoscopy and dye as routine diagnostic procedures in the fertility work-up was evaluated (chapter 3). We concluded that both HSG and laparoscopy cause considerable amount of physical complaints, but the overall judgment of women undergoing the procedures favors laparoscopy.

On the basis of these results, it is difficult to justify the routine performance of hysterosalpingography as an integral part in the basic fertility work-up prior to the invasive laparoscopy.

7.1.3 Alternatives of HSG

Several alternatives can be used to evaluate the tubal function in subfertile women like Chlamydia antibody testing (CAT), hysterosalpingo-contrast-sonography or transvaginal hydrolaparoscopy, of which CAT is most widely used and has been evaluated most extensively.

Chlamydia antibody testing (CAT)

The diagnostic properties of CAT and HSG compared to laparoscopy and dye is well established and showed that CAT performs just as well or better than HSG, but both have poor performance[10-14]. However, the prognostic value of both tests in predicting occurrence of pregnancy is also poor[15]. We came to the same conclusion when we evaluated the value of CAT as alternative test for HSG (chapter 4). However, CAT fails to provide information about the severity of tubal pathology. Furthermore, it cannot detect tubal pathology caused by endometriosis or salpingitis due to other micro-organism than *Chlamydia (C) trachomatis*. However, the costs of CAT are low (in the Netherlands, about ten euros)[16] and patient discomfort is negligible.

CAT as screening test to estimate the risk of tubal pathology prior to laparoscopy is preferable to HSG due to its simplicity and limited inconvenience. The predictive value of CAT for tubal pathology is limited, because the presence of *C. trachomatis* IgG antibodies reflects a previous infection, but not a persistent infection. Measuring serological markers of persistence, of which C-reactive protein (CRP) seems promising and may help to identify women at highest risk of tubal pathology[17]. Further evaluation is needed to decide whether serological markers of persistent *C. trachomatis* infections deserve a place in the fertility work-up[18].

Endovaginal ultrasound

HSG and laparoscopy are not mutually exclusive, as apart from testing tubal patency, they provide different information that may be useful in the evaluation of subfertility[19]. Advantages of HSG relative to laparoscopy are that HSG provides information of the uterine cavity and configuration of the mucosal folds. Endovaginal ultrasound is also recommended as assessment of uterine cavity abnormalities[20]. The ultrasound approach has the additional advantage that the procedure is non-invasive, cheap and may offer supplementary information on tubal and ovarian pathology. The configuration of the mucosal folds is only needed if tubal microsurgery (eg salpingostomy) is feasible. Only if tubal microsurgery is considered, there is an option to perform HSG after a laparoscopy. The success of tubal surgery is determined by the degree of mucosal damage rather than the peritubal adhesions visualized at laparoscopy[21]. However, there is no evidence of benefit or disadvantage of tubal

surgery in the management of tubal infertility versus no treatment or alternative treatments[22].

7.1.4 Recruitment difficulties

It was more difficult than anticipated to recruit subjects to the randomized trial. In chapter five the results of recruitment difficulties in a randomized trial involving HSG are described. Inattentive clinicians and patient-associated reasons (like refusal and indecisiveness to enroll) were the most common reasons not to participate in the trial. Appropriate instruction of the study protocol to the participating doctors and adequate information, in layman's terms, to the patients can overcome non-recruitment. In the set up of the trial and during the recruitment, communication and information are the key words.

7.1.5 Conclusion

Chlamydia antibody testing (CAT) and endovaginal ultrasound may replace HSG as screening test prior to laparoscopy. CAT can be performed in stead of HSG to identify women at high risk for tubal pathology. When the Chlamydia serology findings are negative an expectant management for laparoscopic assessment of tubal pathology may be advised. Likewise, when the Chlamydia serology is abnormal, laparoscopic tubal assessment is indicated. For the assessment of uterine cavity abnormalities an endovaginal ultrasound can be performed. If pathology of the uterine cavity is presumed by ultrasound, hysteroscopy could be performed together with the laparoscopy.

In conclusion, HSG can be omitted from the routine investigation in the fertility workup and should only be applied for specific conditions like contraindication for laparoscopy or examining tubal mucosal fold configuration when tubal microsurgery is considered. Further studies are needed to detect better methods and to make clinical decisions about management on the basis of Chlamydia serology results.

7.2 Therapeutic strategies in reproductive medicine: a study on assisted conception

7.2.1 Perinatal outcome after assisted conception

The second subject presented in this thesis is a systematic review based on observational studies that investigated perinatal outcome as a possible adverse effect of assisted conception.

In the preparation time of the systematic review, we celebrated 25 years of assisted reproduction. The worse perinatal outcome after assisted reproduction was explained

by multiple pregnancy, bias and confounding[23]. In our systematic review of controlled studies published between 1985 and 2002, 25 studies were included of which 17 had matched and 8 had non-matched controls (chapter six). We concluded that singleton pregnancies conceived after assisted conception have significantly worse perinatal outcomes than non-assisted singleton pregnancies (Table 1). A few months later Dr. Mary Croughan's group from California, USA, published a systematic review with an identical conclusion, using by and large the same studies as we have included, but focused only on IVF[24]. In both reviews the same control group (naturally conceived singletons) and definitions were used, and controlled for mostly maternal age and parity. Caesarean rate and admission to a neonatal intensive care unit, probably consequences of the perinatal problems were also significantly increased.

Table 1. Comparison of singletons conceived after assisted conception and spontaneously conceived singletons based on data of Helmerhorst *et al*, BMJ 2004 and Jackson *et al*, Obstet Gynecol 2004

		Helmerhorst		Jackson	
		OR	95%CI	OR	95%CI
LBW	<2500g	1.7	1.5-1.9	1.8	1.4-2.2
VLBW	<1500g	3.0	2.1-4.4	2.7	2.3-3.1
Preterm	<37wks	2.0	1.8-2.3	2.0	1.7-2.2
SGA	$<P_{10}$	1.4	1.2-1.7	1.6	1.3-2.0
PNM	†	1.7	1.1-2.6	2.2	1.6-3.0

Can we trust both reviews? Or in other words is the conclusion true? Observational studies are considered to suffer heavily from bias in contrast to randomised controlled trials. However, the widespread use of hierarchy of evidence that grade research studies according to their quality is constructed for intervention studies, not for identifying infrequently occurring (unexpected) adverse effects[25]. In the situation on assisted conception we are discussing, it is not only unethical to conduct a randomized controlled trial (RCT), even if a RCT could be realized; the daunting number of participants that should be included was unrealistic to finish the trial adequately. Reasons that the conclusion of both reviews might be true, are that the association is consistent in all separate studies, that a dose response gradient is appreciated and that the association is observed in different populations, whereas high quality studies published higher risks. Treatment bias (preterm birth due to obstetric intervention) was not identified by Jackson *et al*.[24].

7.2.2 Possible factors related with adverse perinatal outcome

Is the association between assisted reproductive treatments (ART) and worse perinatal outcome only valid for the IVF technique? In a retrospective cohort study, using data from the regional registry of all hospital deliveries in the Dutch-speaking part of Belgium during the period from January 1993 until December 2003, Ombelet *et al*.[26] compared 12 021 singletons conceived after ovarian stimulation **without** IVF and 12 021 naturally conceived singletons matched for maternal age, parity, fetal sex and year of birth. Their results also showed a significantly higher incidence of prematurity (<32 and <37 weeks) as well as low and very low birth weight among singletons conceived after ovarian stimulation **without** IVF compared to naturally conceived singletons (Table 2).

Table 2. Comparison of singletons conceived after ovarian stimulation without IVF and spontaneously conceived singletons (Ombelet *et al*, Hum Reprod 2006)

		OR	95%CI
LBW	<2500g	1.9	1.7-2.1
VLBW	<1500g	3.2	2.3-4.5
Preterm	<37wks	1.9	1.7-2.1
Very preterm	<32wks	3.3	2.3-4.6

Two matched controlled studies compared perinatal outcome between IVF and IUI singletons and no differences between both groups was observed[27,28]. We may conclude that singleton pregnancies after ovarian stimulation (with or without ART procedure) are associated with worse perinatal outcome. So the ART procedure itself does not attribute to the adverse perinatal outcome of ART pregnancies.

Specific characteristics of an infertility population, such as higher maternal age and lower parity, are well known obstetric risk factors, so that it could *a priori* be expected that pregnancy outcome is poor in this group. Apart from these two confounders, couples with assisted pregnancies differ from couples who naturally conceive in socioeconomic status and smoking, while fetal sex and date of delivery might influence the perinatal outcome[23,29-31]. The matched controlled studies used in both reviews used these additional confounding factors in the studies of Dhont, Koivurova, Koudstaal, Maman, Reubinoff and Verlaenen. These factors seem unable to explain the difference in perinatal outcome.

Whether subfertility in itself may influence the association between assisted conception and worse perinatal outcome is unclear[32]. In order to investigate whether

subfertility explains poor perinatal outcome after assisted conception, Kapiteijn *et al*.[33] (using data from a nation-wide historical cohort of 25 152 women treated for subfertility in the Netherlands between 1980 and 1995) compared perinatal outcome of singletons conceived after controlled ovarian hyperstimulation and IVF (IVF+; n = 2239) with perinatal outcome in subfertile women who conceived spontaneously (subfertile controls without treatment; n = 6343). The subjects were matched for maternal age, BMI, race, education, primiparity, smoking, diabetes, sex infant. Singleton IVF+ pregnancies have significantly worse perinatal outcomes than those of spontaneously conceived pregnancies in subfertile women (table 3). The poor perinatal outcome in this population could not be explained by subfertility and suggests that other factors may be important in the known association between assisted conception and poor perinatal outcome.

Table 3. Comparison of singletons conceived after ovarian stimulation with IVF and subfertile controls without treatment (Kapiteijn *et al*, Hum Reprod 2006)

		OR_{adj}	95%CI
LBW	1500-2500g	1.7	1.4-2.0
VLBW	<1500g	2.7	1.8-4.1
Preterm	32-37wks	1.6	1.3-1.9
Very preterm	<32wks	2.2	1.4-3.3

7.2.3. Clinical implications of worse perinatal outcome

What are the clinical implications of worse perinatal outcome after assisted conception? That depends on the prevalence of ART (IVF) children and the effects of low birth weight in later life. Babies with low birth weight are prone to develop cardiovascular disease in later life , i.e. hypertension, obesity and diabetes. This 'fetal origins of adult disease' or 'Barker hypothesis' suggests that several of the major diseases of later life, including coronary heart disease, stroke and cardiovascular death, originate in impaired intrauterine growth and development[34]. Furthermore, Smith *et al*.[35] found that mothers, who once gave birth to thin babies, have a higher risk of developing ischemic heart disease later in life. Because the risk of cardiovascular disease is known to increase with an increasing number of risk factors, women with a low birth weight should try to avoid acquired risk factors, like smoking and obesity. In addition, extra attention should be given in detecting diabetes or hypertension at a later stage in life and in detecting exaggerated growth during childhood, as these individuals seem most prone to develop cardiovascular disease later on in life.

7.2.4. Implications for the scientist

One of the criteria to determine causation of an association between two epidemiologically found phenomena formulated by Sir Austin Bradford Hill[36] is to identify a biological mechanism behind the association. In an attempt to find a biological plausible explanation for the association between ovarian stimulation and low birth weight and to find on which level ovarian stimulation act its negative effect, Sibug *et al.*[37-39] started form the observation that birth weight of singletons conceived after implanting a cryopreserved embryo is significantly **higher** than birth weight after a **fresh** embryo transfer[40,41]. Predominantly, an cryopreserved-thawed embryo transfer occurs in a natural cycle, thus without direct effect of ovarian stimulation, in contrast to the preceded cycle of IVF where ovarian stimulation is needed for producing sufficient oocytes in order to make selection of them.

One might hypothesize that ovarian hyperstimulation *prior to* IVF might negatively influence uterine receptivity, meaning that ovarian stimulation is associated with low birth weight[42] and might be considered as a negative effect ("stress factor ") on the endometrium[43] and as such on the potential of embryonic implantation. Implantation implies an interaction between the human blastocyt and maternal endometrium. Blastocyts implantation is intimately associated with vascular permeability and angiogenesis. The vascular endothelial growth factor (VEGF) is a potent stimulator of blastocyt implantation. Ovarian stimulation with gonadotrophins to adult female CD1 mice lowered the expression of VEGF and its receptors in embryo implantation sites (37,38). Furthermore, embryo implantation was delayed and the area size of the embryo implantation site was reduced[37]. Two types of gonadotrophins were used (recombinant gonadotrophins and urinary gonadotrophins) as ovarian stimulation in adult female CD1 mice. Urinary gonadotrophins (urinary FSH and urinary hCG) but not recombinant gonadotrophins adversely affect the implantation process[38]. However, caution must be taken in interpreting these results and extrapolating them to the human situation.

7.2.5. Alternative explanations for worse perinatal outcome

To reduce multiple and twin pregnancies, transfer of one embryo (single-embryo transfer (SET)) is offered in selected patient groups. Because the population of patients receiving SET resulting in singleton pregnancies is different from the one that would have become pregnant (with a singleton) before SET, de Neubourg *et al.*[44] prospectively investigated whether the obstetrical outcome of singleton pregnancies after SET differed from spontaneously conceived singletons. They concluded that the incidence of preterm birth was still worse and significantly higher in SET singletons versus spontaneous singletons although the birthweight of these 'preterm born' SET

singletons was not lower than for the spontaneous singletons.

Another plausible explanation for the poorer outcome in IVF singletons is that a sizeable number of IVF singletons origin from twin gestations, where spontaneous reduction occurred, leaving the surviving co-twin with an increased risk of adverse outcome. Recent studies reported that the survivors of a vanishing co-twin are contributing to a higher risk of adverse obstetric outcomes in IVF singletons[45-47].

7.2.6. Conclusion and recommendations

Whatever the explanation may be, singletons conceived after assisted conception have worse perinatal outcome than naturally conceived singletons, but this is substantially less so for twins. The major reason that the outcomes among twins are reaching unity may be found in the assumption that the assisted conception factor among twins is apparently overwhelmed by the multiple conception problems: a comparison between bad and bad remains equal. But this is poor consolation for the much greater risks of twin pregnancy overall (short- and long term outcome measures). Thus the adverse outcome carried in twins is still one of the main concerns of ART.

Our conclusion draw attention to three challenges. Firstly, emphasis needs to shift, more than it has already, from achieving pregnancy to achieving a successful outcome. Secondly, it may be timely to consider any multiple pregnancy after assisted conception as a failure of that technology to achieve what it ought to achieve. Thirdly, there is a need to narrow the gap in perinatal outcome between assisted and other singleton pregnancies. This may also enhance understanding of how gestational age and birth weight interact with each other.

References

1. Helmerhorst FM, Oei SG, Bloemenkamp KW and Keirse MJHC. Consistency and variation in fertility investigations in Europe. Hum Reprod 1995;10:2027-2030.
2. Mol BW, Collins JA, Van Der Veen F and Bossuyt PM. Cost-effectiveness of hysterosalpingography, laparoscopy, and Chlamydia antibody testing in subfertile couples. Fertil Steril 2001;75:571-580.
3. Rowe PJ, Comhaire FH, Hargreave TB and Mellows HJ. WHO manual for the standardized investigation and the diagnosis of the infertile couple. Cambridge University Press;1993,page 4.
4. Swart P, Mol BW, van der Veen F, van Beurden M, Redekop WK and Bossuyt PM. The accuracy of hysterosalpingography in the diagnosis of tubal pathology: a meta-analysis. Fertil Steril 1995;64:486-491.
5. Johnson N, Vandekerckhove P, Watson A, Lilford R, Harada T, Hughes E. Tubal flushing for subfertility. Cochrane Database Syst Rev 2005;2:CD003718.
6. Coppus SF, Opmeer BC, van der Veen F, Bossuyt PM, Mol BW. Routine use of hysterosalpingography prior to diagnostic laparoscopy in the fertility workup. Hum Reprod 2006;21:2725-2726.
7. Haynes B. Can it work? Does it work? Is it worth it? BMJ 1999;319:652–653.
8. Kopitzke EJ, Berg BJ, Wilson JF, Owens. Physical and emotional stress associated with components of the infertility investigation: perspectives of professionals and patients. Fertil Steril 1991;55:1137-1143.
9. Eimers JM, Omtzigt AM, Vogelzang ET, van Ommen R, Habbema JDF, te Velde ER. Physical complaints and emotional stress related to routine diagnostic procedures of the fertility investigation. J Psychosom Obstet Gynaecol 1997;18:31-35.
10. Dabekausen YAJM, Evers JLH, Land JA, Stals FS. *Chlamydia trachomatis* antibody testing is more accurate than hysterosalpingography in predicting tubal factor infertility. Fertil Steril 1994;61:833-837.
11. Meikle SF, Zhang X, Marine WM, Calonge BN, Hamman RF, Betz G. Chlamydia trachomatis antibody titers and hysterosalpingography in predicting tubal disease in infertility patients. Fertil Steril 1994;62:305-312
12. Mol BWJ, Dijkman B, Wertheim P, Lijmer J, van der Veen F, Bossuyt PMM. The accuracy of serum chlamydial antibodies in the diagnosis of tubal pathology: a meta-analysis. Fertil Steril 1997;67:1031-1037.
13. Thomas K, Couglin L, Mannion PT, Haddad NG. The value of *Chlamydia trachomatis* antibody testing as part of routine infertility investigation. Hum Reprod 2000;15:1079-1082.
14. Veenemans LM, Linden PJ. The value of *Chlamydia trachomatis* antibody testing in predicting tubal factor infertility. Hum Reprod 2002;17:695-698.
15. Karande VC, Pratt DE, Rabin DS, Gleicher N. The limited value of hysterosalpingography in assessing tubal status and fertility potential. Fertil Steril 1995;63:1167-1171.
16. Fiddelers AA, Land JA, Voss G, Kessels AGH, Severens JL. Cost-effectiveness of Chlamydia antibody tests in subfertile women. Hum Reprod 2005;20:425-432.

17. Den Hartog JE, Land JA, Stassen FRM, Kessels AGH, Bruggeman CA. Serological markers of persistent C.trachomatis infections in women with tubal factor subfertility. Hum Reprod 2005;20:986-990.
18. Den Hartog JE, Morré SA, Land JA. *Chlamydia trachomatis*-associated tubal factor subfertility: immunogenetic aspects and serological screening. Hum Reprod Update 2006;2:719-730.
19. National Institute for Clinical Excellence, NHS. Fertility: Assessment and Treatment for People with Fertility Problems-Full Guideline. London 2004, RCOG Press.
20. Soares SR, Barbosa dos Reis MM, Camargos AF. Diagnostic accuracy of sonohysterosalpingography, transvaginal sonography, and hysterosalpingography in patients with uterine cavity diseases. Fertil Steril 2000;72:406-411.
21. Vasquez G, Boeckx W, Brosens I. Prospective study of tubal mucosal lesions and fertility in hydrosalpinges. Hum Reprod 1995;10:1075–1078.
22. Ahmad G, Watson A, Vandekerckhove P, Lilford R. Techniques for pelvic surgery in subfertility. *Cochrane Database of Systematic Reviews* 2006, Issue 2.
23. Australian In Vitro Fertilisation Collaborative Group. High incidence of preterm births and early losses in pregnancy after in vitro fertilization. BMJ 1985;291:1160-1163.
24. Jackson RA, Gibson KA, Wu YW, Croughan MS. Perinatal outcomes in singletons following in vitro fertilization: a meta-analysis. Obstet Gynecol 2004;103:551-563.
25. Glasziou P, Vandenbroucke JP, Chalmers I. Assessing the quality of research. BMJ 2004;328:39-41.
26. Ombelet W, Martgens G, de Sutter P, Gerris J, Bosmans E, Ruyssinck G, Defoort P, Molenberghs G, Gyselaers W. Perinatal outcome of 12,021 singleton and 3108 twin births after non-IVF-assisted reproduction: a cohort study. Hum Reprod 2006;21:1025-1032.
27. Nuojua-Huttunen S, Gissler M, Martikainen H, Tuomivaara L. Obstetric and perinatal outcome of pregnancies after intrauterine insemination. Hum Reprod 1999;14:2110-2115.
28. De Sutter P, Veldeman L, Kok P, Szymczak N, Van der Elst J, Dhont M. Comparison of outcome of pregnancy after intra-uterine insemination (IUI) and IVF. Hum Reprod 2005;20:1642-1646.
29. Dhont M, De Sutter P, Ruyssinck G, Martens G, Bekaert A. Perinatal outcome of pregnancies after assisted reproduction: a case-control study. Am J Obstet Gynecol 1999;181:688-695.
30. Tallo CP, Vohr B, Oh W, Rubin LP, Seifer DB, Haning RV Jr. Maternal and neonatal morbidity associated with in vitro fertilization. J Pediatr 1995;127:794-800
31. Lintsen AM, Pasker-de Jong PC, de Boer EJ, Burger CW, Jansen CA, Braat DD, van Leeuwen FE. Effects of subfertility cause, smoking and body weight on the success rate of IVF. Hum Reprod 2005;20:1867-1875.
32. Pandian Z, Bhattacharya S, Templeton A. Review of unexplained infertility and obstetric outcome: a 10 year review. Hum Reprod 2001;16:2593-2597.
33. Kapiteijn K, de Bruijn CS, de Boer E, de Craen AJM, Burger CW, van Leeuwen FE, Helmerhorst FM. Does subfertility explain the risk of poor perinatal outcome after IVF and ovarian hyperstimulation? Hum Reprod 2006;21:3228-3234.

34. Barker DJ. Fetal origins of coronary heart disease. BMJ 1995;311:171-174.
35. Smith GC, Pell JP, Walsh D: Pregnancy complications and maternal risk of ischaemic heart disease: a retrospective cohort study of 129,290 births. Lancet 2001;357:2002-2006.
36. Hill AB. Principles of medical statistics. 9th edition Lancet, London 1971, 313.
37. Sibug RM, Helmerhorst FM, Tijssen AM, Gonadotrophin stimulation reduces VEGF(120) expression in the mouse uterus during the peri-implantation period. Hum Reprod 2002;17:1643-1648.
38. Sibug RM, de Koning J, Tijssen AM, de Kloet ER, de Koning J. Urinary gonadotrophins but not recombinant gonadotrophins reduce expression of $VEGF_{120}$ and its receptors flt-1 and flk-1 in the mouse uterus during the peri-implantation period. Hum Reprod 2005;20:649-656.
39. Sibug RM, Datson N, Tijssen AM, Morsink M, de Koning J, de Kloet ER, Helmerhorst FM. Effects of urinary and recombinant gonadotrophins on gene expression profiles during the murine peri-implantation period. Hum Reprod. 2007;22:75-82.
40. Wennerholm WB. Cryopreservation of embryos and oocytes: obstetric outcome and health in children. Hum Reprod 2000;15:18-25.
41. Wang YA, Sullivan EA, Black D, Eng DJM, Bryant J, Chapman M. Preterm birth and low birth weight after assisted reproductive technology-related pregnancy in Australia between 1996 and 2000. Fertil Steril 2005;83:1650-1658.
42. Olivennes F, Rufat P, Andre B, Pourade A, Quiros MC, Frydman R. The increased risk of complication observed in singleton pregnancies resulting from in-vitro fertilization (IVF) does not seem to be related to the IVF method itself. Hum Reprod 1993;8:1297-1300.
43. Bourgain C, Devroey P. The endometrium in stimulated cycles for IVF. Hum Reprod Update 2003;9:515-522.
44. De Neubourg D, Gerris J, Mangelschots K, van Royen E, Vercruyssen M, Steylemans A, Elseviers M. The obstetrical and neonatal outcome of babies born after single-embryo transfer in IVF/ICSI compares favourably to spontaneously conceived babies. Hum Reprod 2006;21:1041-1046.
45. Verstraelen H, Goetgeluk S, Derom C, Vansteelandt S, Derom R, Goetghebeur E, Temmerman M. Preterm birth in twins after subfertility treatment: population based cohort study.BMJ 2005;331:1173.
46. Pinborg A, Lidegaard O, Andersen AN. The vanishing twin: a major determinant of infant outcome in IVF singleton births. Br J Hosp Med (Lond) 2006;67:417-420.
47. La Sala GB, Nicoli A, Villani MT, Pescarini M, Gallinelli A, Blickstein I. Outcome of 518 salvage oocyte-cryopreservation cycles performed as a routine procedure in an in vitro fertilization program. Fertil Steril 2006;86:1423-1427.

Chapter

8

Summary & Samenvatting

8.1 Summary

In this thesis two studies are presented dealing with diagnostic and therapeutic strategies in reproductive medicine. The first subject is to investigate whether hysterosalpingography (HSG) prior to laparoscopy and dye should be maintained as a diagnostic procedure in the routine fertility work-up. The second subject focuses on therapy in reproductive medicine with the clinically relevant question whether adverse perinatal outcome is related to assisted conception in both singleton and twin pregnancies.

Chapter 1, provides an historical overview of the literature regarding evidence in reproductive medicine.

Because the usefulness of HSG was uncertain, a multicenter randomised controlled trial comparing fertility work-up with or without HSG to assess the value of HSG prior to laparoscopy and dye in a routine clinical setting was carried out. A pragmatic effectiveness trial was designed to evaluate whether two commonly used diagnostic strategies (HSG and laparoscopy & dye) have a different impact on cumulative pregnancy rate of subfertile women (**chapter 2**). A similar cumulative pregnancy rate at 18 months in the intervention (HSG) group (49.1%) [95% confidence interval (CI) 41.6 to 56.6] and the control (non-HSG) group (50.3%) (95% CI 42.8 to 57.8) was found. We concluded that routine use of HSG at an early stage in the fertility work-up prior to laparoscopy and dye does not influence cumulative pregnancy rate compared to routine use of laparoscopy without HSG.

Many women experience HSG as a painful procedure, although laparoscopy is invasive and generally needs general anaesthesia. In an additional analysis of subfertile women nested in the above mentioned randomised controlled trial, pain and patient satisfaction of women undergoing an HSG or laparoscopy and dye as routine diagnostic procedures in the fertility work-up was evaluated and presented in **chapter 3**. The results showed that the HSG procedure was experienced as more painful while laparoscopy and dye was related with more post-procedure complaints. The overall judgment of women undergoing the procedures favours laparoscopy. We concluded that both HSG and laparoscopy cause considerable amount of physical complaints, but the overall judgment of women undergoing the procedures favours laparoscopy.

Chlamydia antibody testing (CAT) is an alternative of HSG to evaluate the tubal function in subfertile women. Nested in the randomised controlled trial, we assessed

in **chapter 4** Chlamydia trachomatis specific IgG antibody testing (CAT) and HSG as test for the likelihood of tubal damage or occurrence of pregnancy prior to laparoscopy and dye. The diagnostic accuracy of both Chlamydia trachomatis specific IgG antibody testing and HSG are comparable but showed poor performance. The prognostic value of both tests in predicting occurrence of pregnancy was also poor. We concluded that CAT as screening test to estimate the risk of tubal pathology prior to laparoscopy is preferable to HSG due to its simplicity and limited inconvenience.

It was more difficult than anticipated to recruit subjects to a randomised controlled trial with or without the performance of HSG. In **chapter 5**, the reasons for non-participation to the randomised trial are described. Inattentive clinicians and patient-associated reasons (like refusal and indecisiveness to enroll) were the most common reasons not to participate in the trial. Appropriate instruction of the study protocol to the participating doctors and adequate information, in layman's terms, to the patients can overcome non-recruitment. In the set up of the trial and during the recruitment, communication and information are the key words.

Chapter 6 describes the results of a systematic review of 25 controlled studies to examine difference in perinatal outcome between assisted and natural conceptions and whether they apply to both singleton and twin pregnancies. We concluded that singleton pregnancies conceived after assisted conception have significantly worse perinatal outcomes than naturally conceived singletons, but this is substantially less so for twins.

In **chapter 7**, the results of the two studies are summarized and discussed in a broader perspective.
In short, the results of the randomised trial involving HSG showed that HSG can be omitted from the routine investigation in the fertility work-up and should only be applied for specific conditions like contraindication for laparoscopy or examining tubal mucosal fold configuration when tubal microsurgery is considered. Chlamydia antibody testing (CAT) and endovaginal ultrasound may replace HSG as screening test prior to laparoscopy. CAT can be performed in stead of HSG to identify women at high risk for tubal pathology. When the Chlamydia serology findings are negative an expectant management for laparoscopic assessment of tubal pathology may be advised. Likewise, when the Chlamydia serology is abnormal, laparoscopic tubal assessment is indicated. For the assessment of uterine cavity abnormality, an endovaginal ultrasound can be performed followed by hysteroscopy and laparoscopy when pathology of the uterine cavity is presumed .

The major reason that the outcomes among twins are reaching unity may be found in the assumption that the assisted conception factor among twins is apparently overwhelmed by the multiple conception problems: a comparison between bad and bad remains equal. But this is poor consolation for the much greater risks of twin pregnancy overall (short- and long term outcome measures). Thus the adverse outcome carried in twins is still one of the main concerns of assisted conception. Our conclusion draws attention to three challenges. Firstly, emphasis needs to shift, more than it has already, from achieving pregnancy to achieving a successful outcome. Secondly, it may be timely to consider any multiple pregnancy after assisted conception as a failure of that technology achieve what it ought to achieve. Thirdly, there is a need to narrow the gap in perinatal outcome between assisted and other singleton pregnancies. This may also enhance understanding of how gestational age and birth weight interact with each other.

8.2 Samenvatting

In dit proefschrift worden twee studies gepresenteerd die diagnostische en therapeutische strategieën in de reproductieve geneeskunde onderzoeken. Het eerste onderwerp onderzoekt of het verrichten van een hysterosalpingografie (HSG) vóórafgaand aan de gouden standaard laparoscopie met tubatesten moet worden gehandhaafd als een standaard diagnostisch onderzoek in het fertiliteitonderzoek. Het tweede onderwerp over behandeling in de reproductieve geneeskunde gaat in op de klinisch relevante vraag of een ongunstige perinatale uitkomst gerelateerd is aan geassisteerde conceptie bij zowel eenling als tweeling zwangerschappen.

Hoofdstuk 1, verschaft een historisch overzicht van de literatuur over bewijs in de reproductieve geneeskunde.

Omdat het nut van HSG twijfelachtig was, werd er een multicenter gerandomiseerd gecontroleerde studie opgezet die een vergelijking maakt tussen fertiliteitonderzoeken met of zonder het verrichten van een HSG om de waarde van HSG vóóraf laparoscopie met tubatesten in een routine klinische setting te bepalen. Een pragmatische effectiviteitstudie werd ontwikkeld om te evalueren of twee vaak toegepaste diagnostische strategieën (HSG en laparoscopie met tubatesten) een verschillende impact hebben op de cumulatieve zwangerschapskans van subfertiele vrouwen (**hoofdstuk 2**). Er was sprake van een gelijke cumulatieve zwangerschapskans bij 18 maanden in zowel de interventie (HSG) groep (49.1%) [95% betrouwbaarheidsinterval (BI) 41.6 tot 56.6] als de controle (niet-HSG) groep (50.3%) [95% BI 42.8 tot 57.8]. Wij concludeerden dat routinematig toepassen van HSG in een vroeg stadium in het fertiliteitonderzoek vóór de laparoscopie met tubatesten, de cumulatieve zwangerschapskans niet beïnvloedt vergeleken met het routinematig toepassen van laparoscopie zonder vooraf HSG.

Veel vrouwen ervaren het HSG als een pijnlijk onderzoek, maar de laparoscopie is een invasieve ingreep dat meestal onder algehele anesthesie plaatsvindt. In een additionele analyse van subfertiele vrouwen die afkomstig waren uit de hierboven beschreven gerandomiseerd gecontroleerde studie, werd de pijn en tevredenheid geëvalueerd en gepresenteerd in **hoofdstuk 3.** Deze groep vrouwen hadden of een HSG of een laparoscopie met tubatesten ondergaan als standaard diagnostisch fertiliteitonderzoek. De resultaten toonden aan dat het HSG onderzoek werd ervaren als een pijnlijker onderzoek, terwijl laparoscopie met tubatesten gerelateerd was aan meer klachten optredend na het onderzoek. De algemene beoordeling van deze vrouwen was ten gunste van de laparoscopie. Wij concludeerden dat zowel HSG als

laparoscopie een aanzienlijke hoeveelheid fysieke klachten veroorzaakt, maar dat de algemene beoordeling van vrouwen die het onderzoek heeft ondergaan ten gunste was van de laparoscopie.

Chlamydia IgG antilichaam titer (CAT) bepaling is een alternatief voor het HSG om de tuba functie te evalueren bij subfertiele vrouwen. Met gegevens afkomstig uit de gerandomiseerd gecontroleerde studie, onderzochten we in **hoofdstuk 4** de waarde van Chlamydia trachomatis specifieke IgG antilichaam titer bepaling (CAT) en HSG als test voor de waarschijnlijkheid van tuba schade of optreden van zwangerschap vóór het verrichten van laparoscopie met tubatesten. De diagnostische nauwkeurigheid van zowel Chlamydia trachomatis specifieke IgG antilichaam titer bepaling en HSG zijn vergelijkbaar maar beide onnauwkeurig. De prognostische waarde van beide testen in voorspellen van het optreden van zwangerschap was even onnauwkeurig. Wij concludeerden dat CAT als test om het risico in te schatten op tuba pathologie vóór de laparoscopie de voorkeur heeft boven HSG vanwege zijn gemak en minder invasiviteit.

Het was moeilijker dan vooraf was geschat om patiënten te rekruteren voor een gerandomiseerd gecontroleerde studie met of zonder het verrichten van een HSG. In **hoofdstuk 5**, worden de redenen van niet-deelname aan deze gerandomiseerde studie beschreven. Onoplettende artsen en patiënt-geassocieerde redenen (zoals weigering en besluiteloosheid om mee te doen) waren de meest voorkomende redenen om niet deel te nemen aan de studie. Het geven van goede instructies over het studieprotocol aan de deelnemende artsen en adequate informatie, in lekentaal aan de patiënten, kan het niet-rekruteren ondervangen.

Hoofdstuk 6 beschrijft de resultaten van een systematische review van 25 gecontroleerde studies om het verschil te onderzoeken in perinatale uitkomsten tussen geassisteerde en spontane conceptie en of dit ook geldt voor eenling en tweeling zwangerschappen. Wij concludeerden dat eenling zwangerschappen ontstaan na geassisteerde conceptie significant slechtere perinatale uitkomsten hebben dan na spontane conceptie, maar dat dit minder geldt voor tweeling zwangerschappen.

In **hoofdstuk 7** zijn de resultaten van de twee studies samengevat en bediscussieerd.
Kort samengevat, de resultaten van de gerandomiseerde HSG studie laten zien dat HSG kan worden weggelaten uit het standaard fertiliteitonderzoek en alleen zou moeten worden verricht om specifieke redenen zoals contra-indicatie voor laparoscopie of beoordelen slijmvliespatroon van de tubae als microchirurgie wordt overwogen.

Chlamydia IgG antilichaam titer (CAT) bepaling en echografisch onderzoek kunnen het HSG vervangen als test vóór de laparoscopie. CAT kan in plaats van HSG worden verricht om vrouwen met verhoogd risico voor tubapathologie te identificeren. Als de Chlamydia serologie uitslagen negatief zijn, dan kan er een afwachtend beleid worden gevoerd ten aanzien van laparoscopische beoordeling voor tubapathologie. Evenzo als de Chlamydia serologie abnormaal is, dan is directe laparoscopische beoordeling van de tubae geïndiceerd. Voor de beoordeling van afwijkingen aan het cavum uteri, kan een echografisch onderzoek worden uitgevoerd gevolgd door hysteroscopie indien er afwijkingen aan het cavum uteri worden vermoed.

De belangrijkste reden dat de resultaten tussen tweelingen een gelijkheid bereiken zou gevonden kunnen worden in de veronderstelling dat de geassisteerde conceptie factor tussen tweelingen schijnbaar overstelpt wordt door het meerling conceptie probleem: een vergelijking tussen slecht en slecht blijft gelijk. Maar dit is schrale troost voor de veel grotere risico's van tweeling zwangerschap in algemene zin (korte- en lange termijn uitkomstmaten). Dus de ongunstige uitkomst bij tweelingen is nog steeds een van de grootste zorgen van geassisteerde conceptie. Onze conclusie trekt de aandacht naar drie uitdagingen. Als eerste moet de nadruk verschuiven, meer dan het thans is, van volbrengen tot een zwangerschap naar volbrengen tot een succesvolle uitkomst. Ten tweede wordt het tijd om elke meerling zwangerschap na geassisteerde conceptie te zien als een gebrek van de technologie. Als derde is het nodig om de ruimte tussen de perinatale uitkomst tussen geassisteerde en spontane eenling zwangerschappen te verkleinen. Dit zou dan ook meer duidelijkheid kunnen geven over hoe zwangerschapsduur en geboortegewicht aan elkaar gerelateerd zijn.

Addendum

VRAGENLIJST vooraf HSG of laparoscopie

Geachte mevrouw,

Wij vragen u deze vragenlijst voor aanvang van het HSG of laparoscopie in te vullen. Sommige van onderstaande woorden omschrijven pijngevoelens, die u heeft gehad tijdens de EERSTE DAG van uw laatste menstruatie.
Van de 20 rijtjes woorden mag u er één uitkiezen die het beste bij uw pijn tijdens de eerste dag van de laatste menstruatie past. Als er in een rijtje geen enkel woord staat dat op uw pijn van toepassing is, ga dan verder met het volgende rijtje.
Het tweede deel van de vragenlijst bestaat uit een lijn met links de woorden "Geen enkele" en rechts de woorden "Maximaal voorstelbare". De bedoeling is dat u op deze lijn één kruisje zet om aan te geven hoe pijnlijk de eerste dag van uw laatste menstruatie was.

Hartelijk bedankt voor het invullen van de vragenlijst.

Uw pijn tijdens de eerste dag van uw laatste menstruatie is:
(juiste antwoord omcirkelen)

1. kloppend - bonzend - barstend
2. opflikkerend - flitsend - schietend
3. prikkend - stekend - doorborend
4. scherp - snijdend - messcherp
5. drukkend - knellend - snoerend
6. trekkend - splijtend - scheurend
7. branderig - brandend - vlammend
8. broeiend - gloeiend - verschroeiend
9. koud - ijskoud - vriezend
10. tintelend - jeukend - elektrisch
11. stijf - strak - krampend
12. zeurend - knagend - hardnekkig
13. vermoeiend - afmattend - uitputtend
14. chagrijnig - deprimerend - ziekmakend
15. gespannen - benauwend - verstikkend
16. verontrustend - beangstigend - angstaanjagend
17. treiterend - kwellend - martelend
18. licht - matig - erg - enorm
19. draaglijk - hinderlijk - ontzettend - onhoudbaar
20. vervelend - ellendig - vreselijk – afgrijselijk

Randomisatie nummer:
Datum van invullen van deze vragenlijst:

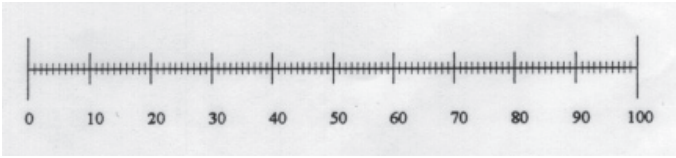
Wij willen u vragen om op onderstaande meetschaal aan te geven hoe pijnlijk u de EERSTE DAG van uw laatste menstruatie heeft ervaren. Op de meetschaal betekent '100' **ondraaglijke-maximale-pijn** en '0' betekent **geen pijn**. De bedoeling is dat u op deze lijn één kruisje tekent dat volgens u de pijn aangeeft die u heeft ervaren.

1.

Hoe is de **gemiddelde pijn** die u op de eerste dag van de laatste menstruatie heeft ervaren?

Gemiddelde pijn zoals u die op de eerste dag van de laatste menstruatie heeft ervaren

Geen enkele



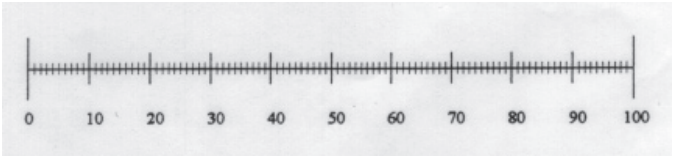
Maximaal voorstelbare

2.

Hoe is de **maximale pijn** die u op de eerste dag van de laatste menstruatie heeft ervaren?

Maximale pijn zoals u die op de eerste dag van de laatste menstruatie heeft ervaren

Geen enkele



Maximaal voorstelbare

VRAGENLIJST na HSG of laparoscopie

Geachte mevrouw,

Wij vragen u deze vragenlijst in te vullen nadat u een HSG of laparoscopie heeft ondergaan.
Het eerste deel van de vragenlijst omschrijft pijngevoelens die u heeft gehad direkt na het onderzoek. Van de 20 rijtjes woorden mag u er één uitkiezen die het beste bij uw pijn past. Als er in een rijtje geen enkel woord staat dat op uw pijn van toepassing is, ga dan verder met het volgende rijtje.

Het tweede deel van de vragenlijst bestaat uit een lijn met links de woorden "Geen enkele" en rechts de woorden "Maximaal voorstelbare". De bedoeling is dat u op deze lijn één kruisje zet om aan te geven hoe pijnlijk u het onderzoek heeft ervaren.

Hartelijk bedankt voor het invullen van de vragenlijst.

Uw pijn na het onderzoek was:
(juiste antwoord omcirkelen)

1. kloppend - bonzend - barstend
2. opflikkerend - flitsend - schietend
3. prikkend - stekend - doorborend
4. scherp - snijdend - messcherp
5. drukkend - knellend - snoerend
6. trekkend - splijtend - scheurend
7. branderig - brandend - vlammend
8. broeiend - gloeiend - verschroeiend
9. koud - ijskoud - vriezend
10. tintelend - jeukend - elektrisch
11. stijf - strak - krampend
12. zeurend - knagend - hardnekkig
13. vermoeiend - afmattend - uitputtend
14. chagrijnig - deprimerend - ziekmakend
15. gespannen - benauwend - verstikkend
16. verontrustend - beangstigend - angstaanjagend
17. treiterend - kwellend - martelend
18. licht - matig - erg - enorm
19. draaglijk - hinderlijk - ontzettend - onhoudbaar
20. vervelend - ellendig - vreselijk – afgrijselijk

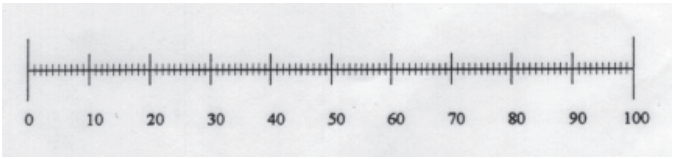
Randomisatie nummer:
Datum van invullen van deze vragenlijst:

Wij willen u vragen om op onderstaande meetschaal aan te geven hoe pijnlijk u het onderzoek heeft ervaren. Op de meetschaal betekent '100' **ondraaglijke-maximale-pijn** en '0' betekent **geen pijn**. De bedoeling is dat u op deze lijn één kruisje tekent dat volgens u de pijn aangeeft die u heeft ervaren.

1.

Hoe is de **gemiddelde pijn** die u na het onderzoek heeft ervaren?

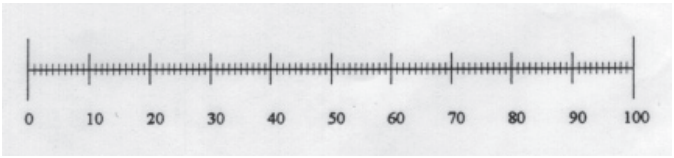
Gemiddelde pijn zoals u die na het onderzoek heeft ervaren

Geen enkele  Maximaal voorstelbare

2.

Hoe is de **maximale pijn** die u na het onderzoek heeft ervaren?

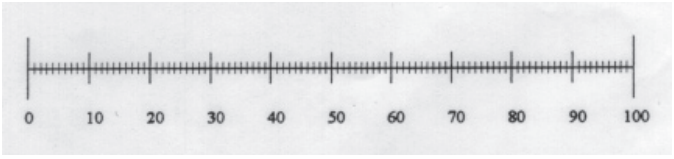
Maximale pijn zoals u die na het onderzoek heeft ervaren

Geen enkele  Maximaal voorstelbare

3.

Hoe is de pijn zoals u die **nu**, na het onderzoek, voelt?

Pijn zoals u die nu, na het onderzoek voelt

Geen enkele  Maximaal voorstelbare

VRAGENLIJST 1 week na HSG of laparoscopie

Geachte mevrouw,

Wij vragen u deze vragenlijst in te vullen 1 week nadat u een HSG of laparoscopie heeft ondergaan. Er worden een aantal vragen gesteld over de pijn waar u last van heeft gehad na het ondergaan van dit onderzoek. Als u alle vragen heeft beantwoord dan kunt u met een toegevoegde antwoordenveloppe de brief terugsturen naar het ziekenhuis.

Hartelijk bedankt voor het invullen van de vragenlijst.

Lees elke zin nauwkeurig door en bepaal in hoeverre u last gehad heeft van deze bijwerkingen. U kunt de vraag beantwoorden door het juiste antwoord te omcirkelen.
1=helemaal niet, 2=een beetje, 3=redelijk wat, 4=veel, 5=heel veel

1.	In hoeverre heeft u voldoende informatie gekregen over het HSG of laparoscopie	**1**	**2**	**3**	**4**	**5**
2.	In hoeverre heeft u last gehad na het onderzoek van:					
	buikpijn	**1**	**2**	**3**	**4**	**5**
	bloedverlies	**1**	**2**	**3**	**4**	**5**
	schouderpijn	**1**	**2**	**3**	**4**	**5**
	misselijkheid	**1**	**2**	**3**	**4**	**5**
	braken	**1**	**2**	**3**	**4**	**5**
	duizeligheid	**1**	**2**	**3**	**4**	**5**
	hoofdpijn	**1**	**2**	**3**	**4**	**5**
	keelpijn	**1**	**2**	**3**	**4**	**5**
	spierpijn	**1**	**2**	**3**	**4**	**5**
	moeheid, algehele malaise	**1**	**2**	**3**	**4**	**5**
3a.	In hoeverre vond u de volgende aspecten van het HSG vervelend:					
	wachttijd onderzoek	**1**	**2**	**3**	**4**	**5**
	inbrengen speculum	**1**	**2**	**3**	**4**	**5**
	doorspuiten eileiders	**1**	**2**	**3**	**4**	**5**
	plaatsen cupje	**1**	**2**	**3**	**4**	**5**
	bloedverlies	**1**	**2**	**3**	**4**	**5**

3b.	In hoeverre vond u de volgende aspecten van de laparoscopie vervelend:					
	wachttijd onderzoek	**1**	**2**	**3**	**4**	**5**
	opname ziekenhuis	**1**	**2**	**3**	**4**	**5**
	Het onder narcose moeten	**1**	**2**	**3**	**4**	**5**
	bloedverlies	**1**	**2**	**3**	**4**	**5**
4.	In hoeverre is het onderzoek u tegengevallen?	**1**	**2**	**3**	**4**	**5**
5.	Heeft u nog pijnstillers moeten gebruiken?	Ja / nee				
6.	Na hoeveel dagen kon u weer normale aktiviteiten verrichten?	1-2-3-4-5->5				
7.	Als u alles vantevoren had geweten, zou u de ingreep opnieuw doen?	Ja / nee				
8.	Zou u de ingreep een vriendin aanraden?	Ja / nee				

Dankwoord

Het is af!! Na meer dan 10 jaar kan ik de zo vaak gestelde vraag “Is het af?” met ja beantwoorden, vol trots schrijf ik dan ook dit dankwoord. Maar dit proefschrift had nooit gerealiseerd kunnen worden zonder de onmisbare hulp en steun van vele mensen. Enkelen wil ik hiervoor in het bijzonder bedanken.

Dr. Trudy Trimbos-Kemper en Erik Hallensleben voor hun hulp bij de uitvoering van het HSG onderzoek.

Prof. Dr. Ir. Anne Stiggelbout voor het maken van de pijnvragenlijsten en hulp bij de opzet van de statistiek voor het HSG onderzoek.

Leonie Louwé voor het beschikbaar stellen van haar werkkamer, de ruimte en de rust gaven mij een duw in de goede richting. Ook de cappuccino en de dropjes bij Wil en Bea om de hoek waren een heerlijke afleiding.

Miranda de Vreede en Marian Bosch-Mackay voor de secretariële ondersteuning.

Thomas Vissers, bibliothecaris van Medisch Centrum Haaglanden, voor alle hulp bij het opzoeken en opvragen van de vele artikelen.

Oud assistenten Noush, Claudia en Caroline waren contactpersoon voor het HSG onderzoek. Ik dank jullie voor het zorgvuldig letten op de inclusie en invullen van de vragenlijsten.

Kim Urgel, destijds keuze-coassistent, heeft veel werk verricht voor het onderzoeken van de non-respons groep, het werd een onderdeel van hoofdstuk 5.

Diane Donker, destijds keuze-coassistent, heeft onder begeleiding van Prof. Dr. Marc Keirse, een grote bijdrage geleverd aan de meta-analyse naar de perinatale uitkomsten van eenlingen en tweelingen na geassisteerde conceptie. Het werd de basis van hoofdstuk 6.

De maatschap Gynaecologie van Medisch Centrum Haaglanden dank ik voor hun onontbeerlijke steun. Alle functies heb ik bij jullie doorlopen, de ruimte die ik hierdoor heb gekregen gaf mij de mogelijkheid om te promoveren.

Mijn maten Doitse, Gerrit, Hans, Job, Leonard, Maria, Mirjam en Taeke wil ik danken voor het volle vertrouwen dat jullie in mij hebben. Ook het “fertiliteitteam” Job, Taeke, Lydia ondersteund door Yda, Anja, Ria, Elly, Lieuwkje, Renuka, Regina en

Anjelien wil ik danken voor de gezellige sfeer en ware teamgeest.

Door de goede opvang thuis door Adèle, Brigitte en Annamarie hoefde ik mij nooit zorgen te maken over de kinderen.

Geerte, eerst was je mijn getuige, nu ben je mijn paranimf. Het is bijzonder om zo´n goede vriendin te hebben! Wat geweldig dat jij nu ook in het noorden woont.

Carla, onze vriendschap begon al in het begin van onze geneeskunde studie. Dank voor alle adviezen die je me op welk gebied dan ook, de afgelopen jaren hebt gegeven.

Pa en mama, zonder jullie was dit proefschrift nooit gekomen. Altijd stonden jullie voor me klaar, ook op de meest onmogelijke tijdstippen. Jullie gaven mij de ambitie en het doorzettingsvermogen om dit proefschrift af te ronden. Ik bof met zulke lieve ouders.

Voor hun niet aflatende belangstelling, begrip en gezelligheid wil ik Willem en Eveline, mijn schoonouders, Joost en Marie en mijn nichtjes Helen en Petra heel erg bedanken. Ik hoop dat ook mijn lieve bijna 100-jarige opa begrijpt dat het proefschrift er is.

Berend en Wouter wil ik bedanken voor al hun geduld, hoe moeilijk het ook voor jullie was om mama achter de computer te laten zitten.

Last but not least (hoe cliché dan ook) wil ik Maarten bedanken. Het valt niet mee om een partner te zijn van een jarenlange promovendus. Dank voor jouw onvoorwaardelijke steun en vertrouwen!

Curriculum Vitae

Denise Perquin werd geboren op 13 december 1968 in Dordrecht. Zij behaalde het VWO diploma in 1988, waarna zij in datzelfde jaar begon aan de studie Geneeskunde aan de Rijksuniversiteit Leiden. In 1989 behaalde ze het propedeutisch examen. Het studiejaar 1992-1993 werd vervolgd aan de Universiteit van Bristol in Engeland via een Exchange Credit Transfer System en Erasmus beurs. Het doctoraal examen werd behaald in 1993, waarna zij begon met de co-schappen. Tijdens het co-schap Gynaecologie en Verloskunde in Medisch Centrum Haaglanden, deed zij een wetenschappelijk onderzoek over de waarde van het HSG in het fertiliteitsonderzoek. Als keuze co-schap breidde zij dit onderzoek uit onder begeleiding van Dr. P.J. Dörr en Prof. Dr. F.M. Helmerhorst. Zij deed hierna een extra keuze co-schap Gynaecologie en Verloskunde in het Kalafong ziekenhuis in Pretoria, Zuid-Afrika, onder supervisie van Prof. R.C. Pattinson. Na terugkomst behaalde zij haar artsexamen in 1995 en was zij van 1995 tot 1998 AGNIO op de afdeling Gynaecologie en Verloskunde in Medisch Centrum Haaglanden. Tijdens dit assistentschap werd een multicenter gerandomiseerd gecontroleerde studie opgezet over de waarde van het HSG in het fertiliteitsonderzoek onder begeleiding van haar promotor en co-promotores; dit onderzoek werd de basis van dit proefschrift. Vervolgens ging zij in 1998 een half jaar naar Flinders Medical Center in Adelaide, Australië, waar ze wetenschappelijk onderzoek deed over de vergelijking van perinatale uitkomsten van tweelingen ontstaan spontaan en na geassisteerde conceptie in Zuid-Australië, onder begeleiding van Prof. Dr. M.J.N.C. Keirse en Prof. Dr. F.M. Helmerhorst. Deze resultaten werden de basis voor het doen van een review over perinatale uitkomsten van eenlingen en tweelingen, een onderdeel van dit proefschrift. In 1999 begon zij met de opleiding tot gynaecoloog in het Leids Universitair Medisch Centrum (opleider: Prof. Dr. H.H.H. Kanhai) en het Medisch Centrum Haaglanden (opleider: Dr P.J. Dörr). In april 2005 volgde inschrijving in het specialistenregister. Vanaf juli 2005 was zij werkzaam als ‘chef de clinique’ in Medisch Centrum Haaglanden. Vanaf 1 januari 2007 is zij staflid Gynaecologie en Verloskunde in Medisch Centrum Leeuwarden.