IMPROVING PREGNANCY CHANCES AND PATIENT EXPERIENCES IN INFERTILITY

KIMMY ROSIELLE

IMPROVING PREGNANCY CHANCES AND PATIENT EXPERIENCES IN INFERTILITY

KIMMY ROSIELLE

Improving pregnancy chances and patient experiences in infertility

Financial support for printing this thesis was kindly provided and supported by the Amsterdam Reproduction & Development (AR&D) research institute, Bridea Medical BV, Chipsoft, Gedeon Richter, Goodlife Pharma, Guerbet and SyncVR.

Provided by thesis specialist Ridderprint, ridderprint.nl

ISBN:	978-94-6483-197-9
DOI:	http://doi.org/10.5463/thesis.139
Cover design & layout:	Marilou Maes, persoonlijkproefschrift.nl
Printing:	Ridderprint

VRIJE UNIVERSITEIT

Improving pregnancy chances and patient experiences in infertility

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. J.J.G. Geurts,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Geneeskunde
op donderdag 22 juni 2023 om 11.45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

door

Kimmy Rosielle
geboren te Doorn

promotoren:	prof.dr. V. Mijatovic prof.dr. C.B. Lambalk
copromotoren:	prof.dr. B.W.J. Mol dr. K. Dreyer
promotiecommissie:	prof.dr. M.A. Oudijk prof.dr. A. Hoek prof.dr. J.S.E. Laven prof.dr. D. de Neubourg dr. J.W. Kallewaard dr. J.H.T.M. van Waesberghe

CONTENTS

Chapter 1	General introduction	7
Chapter 2	How long does the fertility-enhancing effect of hysterosalpingography with oil-based contrast last? *Published in Reproductive BioMedicine Online 2020 doi: 10.1016/j.rbmo.2020.-08.038*	23
Chapter 3	Oil-based versus water-based contrast media for hysterosalpingography in infertile women of advanced age, with ovulation disorders or a high risk for tubal pathology: study protocol of a randomized controlled trial (H2Oil2 study) *Published in BMC Women's Health 2022 doi: 10.1186/s12905-022-01707-z*	41
Chapter 4	Does the effectiveness of IUI in couples with unexplained subfertility depend on their prognosis of natural conception? A replication of the H2Oil study *Published in Human Reproduction Open 2020 doi:10.1093/hropen/hoaa047*	61
Chapter 5	Safety of HSG with oil-based contrast medium: a Systematic review *Published in Reproductive BioMedicine Online 2021 doi: 10.1016/j.rbmo.2021-.03.014*	79
Chapter 6	Virtual Reality for pain relief in hysterosalpingography *Manuscript in preparation*	129
Chapter 7	Visual tubal patency tests for tubal occlusion and hydrosalpinx *Published in Cochrane Database of Systematic Reviews 2022 doi/10.1002/-14651858.CD014968/full*	153
Chapter 8	The impact of the COVID-19 pandemic on infertility patients and endometriosis patients in the Netherlands *Published in reproductive BioMedicine Online 2021 https://doi.org/10.1016/j.rbmo.2021.06.001 1472-6483/*	171
Chapter 9	General discussion and future perspectives	191
Appendix	Summary	208
	Samenvatting	212
	List of authors	216
	List of publications	218
	PhD portfolio	220
	Dankwoord	222
	About the author	227

CHAPTER 1

General introduction

INFERTILITY

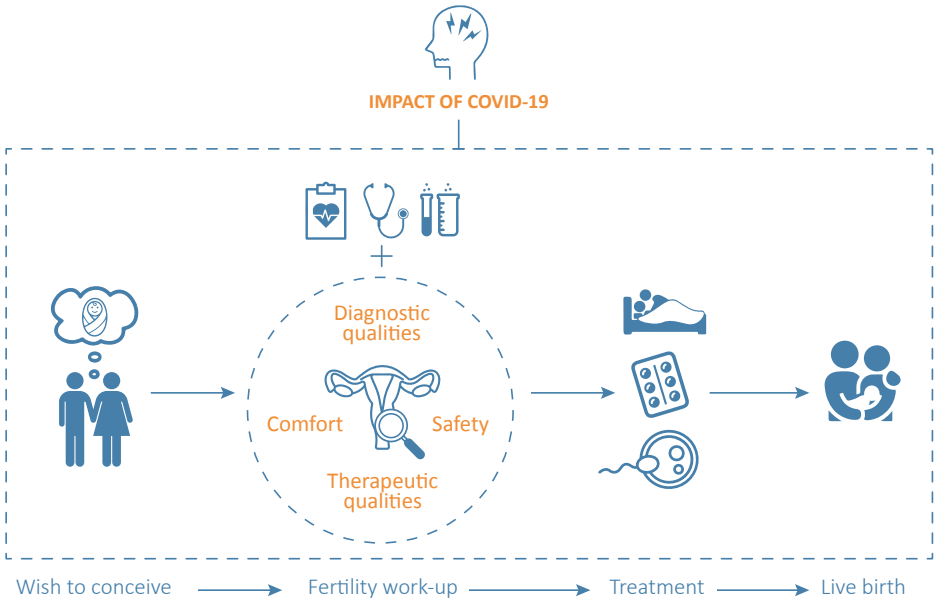
The inability to conceive within 12 months of unprotected intercourse is referred to as infertility, a disease of the female and/or male reproductive system (1, 2). In 2010 the total amount of couples facing infertility was estimated to be 48.5 million couples worldwide (3) which translates to an estimated prevalence of one in six couples trying to conceive (4). More recently, the trend of the prevalence of infertility was modelled based on data from the Global Burden of Disease Study 2017 (5), showing a gradual increase in infertility over time (6). Couples facing infertility can be referred for fertility work-up, a cluster of diagnostic tests investigating the possible causes of infertility in both the female and male partner. The results of the fertility work-up are used to determine the best strategy to fulfil their wish to conceive (7, 8).

Causes of female infertility can be categorized as ovulation disorders, diminished ovarian reserve, uterine, ovarian and/ – or tubal pathology. The respective prevalences of these causes vary depending on the studied population (4, 9). Male factor infertility refers to an impaired number of motile spermatozoa which can have a congenital or acquired cause (10). The diagnosis of 'unexplained infertility' is made when no abnormalities are found in both the female and male partner during the fertility work-up (7, 8).

FERTILITY WORK-UP

The first medical step for couples facing infertility is the fertility work-up (see figure 1). Fertility work-up starts with an extensive medical history of both partners to identify risk factors potentially contributing to infertility such as congenital disorders, acquired diseases or surgical interventions. General blood testing is usually performed in women to check the overall health and to screen for a past infection with *Chlamydia trachomatis* using a *Chlamydia trachomatis* Antibody Test (CAT). An infection with this sexually transmittable pathogen can go unnoticed in women or manifest itself as clinical pelvic inflammatory disease (PID). A positive CAT has been associated with the occurrence of tubal pathology and intra-abdominal adhesions, however, CAT has a low positive predictive value for tubal factor infertility (11, 12). Other events in the medical history that may indicate a risk for tubal pathology are a history of PID, a complicated appendicitis, pelvic surgery, endometriosis and ectopic pregnancy (13). Gynaecological examination combined with transvaginal ultrasound is performed to rule out gynaecological abnormalities such as congenital uterine anomalies, uterine and adnexal pathology, endometriosis and to assess the ovarian reserve.

Figure 1. The journey from wish to conceive to a live born child for infertile couples from left to right.



Couples with a wish to conceive are labelled as 'infertile' after 12 months of unsuccesful trying. They can be referred for infertility work-up to diagnose the type of infertility. Depending on the infertility diagnosis, a treatment plan is proposed. The orange words depict the clinical aspects this thesis will adress.

TUBAL PATENCY TESTING

Tubal patency can be assessed by infusing contrast into the uterine cavity and confirming whether this contrast reaches the intra-abdominal cavity through the Fallopian tubes. Nowadays, there are many factors to take into consideration when determining the most appropriate tubal patency test for women facing infertility. This is represented in the smaller circle of figure 1. The diagnostic quality of a test, the sensitivity and specificity for diagnosing a condition with the given test, is the first factor to take into consideration. Therapeutic qualities refer to any beneficial effect or side-effect of the given test. As with any medical intervention, safety should be considered and lastly, the patient experience should be factored in as well.

Different tests for the assessment of tubal patency were developed at the start of the 20th century. An early method of tubal patency test was developed by Rubin in 1919 (14). He initially inflated the uterus with oxygen, and later carbon dioxide (which reduced the risk of a gas embolus considerably due to the ability of haemoglobin to

absorb carbon dioxide), and expected this gas to escape through the Fallopian tubes in case of patency, inducing a pneumoperitoneum. There are several ways to detect this pneumoperitoneum: the gas gives a painful sensation to the shoulder as it irritates the peritoneal cavity, it is often audible by percussion or auscultation of the thorax and it is visible on X-ray. If the gas inflow showed initial high resistance which then decreased, this also confirmed that at least one of the Fallopian tubes was deemed patent (15). This test was not able to differentiate between unilateral or bilateral tubal patency and caused significant discomfort, and therefore other tubal patency tests were developed.

At the same time, surgical techniques developed greatly across the 20th century. Initially, laparotomic surgery was the only modality available. Laparotomic surgery comes with a high risk of complications, a long recovery time and it can induce intra-abdominal adhesions at the surgical site. Intra-abdominal adhesions are a known risk factor for tubal infertility and should be prevented where possible. The first laparoscopic surgery in humans was performed in 1910 by the Swedish physician Hans Christian Jacobaeus. The technique has only been used for evaluation of tubal patency since the mid 1980's when chromopertubation was introduced. During chromopertubation, a methylene blue is inserted into the uterine cavity transvaginally and flow of this dye from the Fallopian tubes into the peritoneal cavity can be visualized directly. Laparoscopy is still considered as the gold standard test for evaluation of tubal patency (16).

However, laparoscopy is also invasive and requires full anaesthesia. It is therefore usually reserved for women in whom tubal pathology is suspected by less invasive tubal patency tests, or for women in whom an immediate therapeutic intervention during laparoscopy is desired, for instance in case of endometriosis and/or adhesions. As an outpatient alternative to laparoscopy, transvaginal hydrolaparoscopy, also referred to as fertiloscopy, has been introduced, which consists of a transvaginal hydropelviscopy performed under local anaesthesia (17). This approach allows a complete exploration of the pelvis except for the visualization of the vesico-uterine pouch. Apart from tubal patency testing, it allows direct visualization of the tubal mucosa as salpingoscopy is feasible from the vaginal approach, ovarian drilling and limited adhaesiolysis or electrocoagulation of peritoneal endometriosis.

HSG

The first visual tubal patency test, hysterosalpingography (HSG), was performed by Rindfleisch in 1910. During HSG, a radiopaque contrast fluid is injected into the uterus transvaginally. Once the uterus and Fallopian tubes fill with contrast, this can be visualized on X-ray images. Using a suspension of bismuth, a non-toxic heavy metal as a contrast medium, his aim was to outline the uterine cavity in order to detect abnormalities of the

uterus (18). Although this provided good diagnostic images, the suspension remained in the body for a long time and therefore another type of contrast medium was sought to replace bismuth. Lipiodol (Guerbet, Villepinte, France), an iodinated oil-solution, was developed in 1901 as a treatment for goiter, syphilis, cardiovascular problems and other diseases (19). Its first use as contrast medium during HSG was described in 1922. HSG with Lipiodol as oil-based contrast became a common tubal patency test in the investigation of infertility, suspected anomalies in the female genital tract and even to visualize and confirm intrauterine pregnancy in an era where pregnancy tests were lacking. Over the following decades, water soluble contrast media were developed as an alternative to oil-based contrast in HSG. The use of the less viscose water-based contrast media facilitated the implementation of serial HSG, during which a series of X-ray images are taken to evaluate the progression of contrast through the reproductive system. This dynamic approach provides more information than the initial HSG procedure where only one image is obtained hours or even a day after infusion of contrast medium to determine whether there is tubal patency (20). In the 1970's this technique was implemented by tubal surgeons. Before IVF was introduced, tubal surgery was the only option for patients with tubal infertility and serial HSG allowed the surgeon to assess whether surgery could improve fertility. Especially, the visual appearance of water-based contrast made it possible to determine more accurately the preoperative chances of successful tubal reconstructive surgery when compared to the use of oil-based contrast as it is able to show more clearly the mucosal folds in the ampullar part of the tubes.

Fertility enhancing effect of tubal flushing

Since the 1930's it has been reported that Lipiodol and later Lipiodol Ultra Fluid during HSG seemed to have a positive effect on the chance of pregnancy (21, 22). This potential pregnancy enhancing effect of oil-based contrast was the topic of interest in the H2Oil study (23). The H2Oil study was a multicentre randomized controlled trial (RCT) in the Netherlands, investigating the difference in ongoing pregnancies in couples with unexplained or mild male infertility within 6 months after tubal flushing during HSG with oil-based or water-based contrast. The 27 participating hospitals included a total of 1119 women, of which 557 women were randomized for HSG with use of oil-based contrast and 562 for water-based contrast. The ongoing pregnancy rate was significantly higher in the group receiving oil-based contrast (39.7 versus 29.1%, rate ratio (RR) 1.37; 95% confidence interval (CI) 1.16 to 1.61; P<0.001). A recent network meta-analysis included the H2Oil study and showed that the chance of clinical pregnancy within six months was significantly higher after tubal flushing with oil-based contrast than after tubal flushing with water-based contrast (five studies included, OR 1.67, 95% CI 1.38-2.03) (24). Live birth as outcome was only reported in two RCT's (OR 2.18, 95% CI 1.30-3.65) and the certainty of evidence was low. The studies in this network meta-analysis mostly included

couples with unexplained or mild male infertility. Since the publication of this network meta-analysis, a Cochrane systematic review on this topic was published in 2020 which confirmed these findings (25). A Chinese replication study of the original H2Oil study was published in 2022, also showing a similar difference in ongoing pregnancy after the use of oil-based contrast during HSG when compared to water-based contrast (29.1 versus 20.1%, RR 1.44 , 95% CI 1.15-1.81, p=0.001) as well as a shorter time to pregnancy in favour of oil-based contrast (26). Results of large studies on the effect of tubal flushing in couples with other factors of infertility such as ovulation disorders, high risk of tubal pathology and advanced female age are currently lacking.

Potential mechanisms

To obtain a better understanding of which patient groups will benefit from the use of oil-based flushing, it is relevant to unravel the mechanism of action of the fertility enhancing effect. Different hypotheses explaining the mechanism have been formulated (27), placing the points of action in the endometrium (28), the fallopian tube (29), and the peritoneum (30, 31).

The first theory is that the oil-based contrast, derived from poppy seed oil, affects the receptivity of the endometrium through opium receptors present in endometrial cells (32). This may enhance embryo implantation. Another theory hypothesizes that tubal flushing with oil-based contrast flushes accumulated debris and mucous plugs from the proximal part of the otherwise undamaged tubes, enhancing patency and improving the tubal cilia operation (27, 29). A post-hoc analysis of the H2Oil-study showed that the benefit of oil-based contrast is especially visible in women with higher pain scores during HSG, as in women with moderate to severe pain during HSG, the ongoing pregnancy rate was higher following the use of oil-based contrast compared to the use of water-based contrast (33). In women with lower pain scores, there was no difference in ongoing pregnancy rates between oil-based or water-based contrast. These results support the hypothesis that the fertility enhancing effect is associated with a mechanical flushing effect in the fallopian tube. The reported pain is presumed to be caused by an increase in intrauterine pressure prior to dislodgment of pregnancy-hindering debris from the proximal part of otherwise anatomically normal fallopian tubes. One of the several differences in the chemical and physical characteristics between the two types of contrast media might be responsible for the increased flushing potential of oil-based contrast (33). Another possible mechanism taking place in the Fallopian tubes is enhancement of ciliary activity when the oil-based contrast surrounds the cilia and reduces friction. A decrease in friction will improve the function of the cilia, generating flow inside the Fallopian tubes and facilitating transport of gametes (27).

Another theory is modulation of peritoneal macrophage activity by oil-based contrast which affects the conception and implantation mechanism positively (30). In vitro studies have demonstrated that oil-based contrast inhibits phagocytosis of macrophages in humans and rats, perchance phagocytosis of sperm (30, 31). A study by Izumi et al. confirms that oil-based contrast is incorporated in dendritic cells in the peritoneal cavity (34). By promoting more mature dendritic cells, altering cytokine and chemokines profiles in dendritic cells and increasing the number of T-cells, this modulates the immunological environment in the peritoneal cavity. Mature dendritic cells are likely to show decreased activity in sperm phagocytosis. Furthermore, the cytokine IL-10 in the endometrium is upregulated by oil-based contrast while a low presence of IL-10 has been related to repeated implantation failure. Together these may contribute to the fertility enhancing effect of HSG with oil contrast (34).

In addition to the above mentioned hypotheses, it is possible that tubal flushing during HSG improves sperm transport through the cervix (35) as HSG requires manipulation and instrumentation to the cervix. Further research, both fundamental and clinical, is needed in order to provide more insight into the true mechanism(s), duration and dynamics of action of oil-based contrast.

Safety

Shortly after the first reports on the use of HSGs for the treatment of infertility, publications reporting on complications of HSG with Lipiodol emerged. The main severe complication was flow of contrast into the myometrial vessels, potentially leading to contrast entering the venous or lymphatic system (36). This complication is referred to as intravasation and may cause allergic reactions and oil embolisms with potentially harmful or even fatal results (37). Intravasation was attributed to the high viscosity and hydrophobic qualities of Lipiodol and therefore less viscose contrast media were developed, resulting in Lipiodol Ultra Fluid and various types of water-based contrast media (19).

In a recent survey study, intravasation of oil-based contrast during HSG was estimated to occur in 5% of the cases in the Netherlands (38). An overview of the complete literature on the differences in complication rates between HSG with oil-based contrast and water-based contrast is lacking.

Discomfort

Aside from the complication risk, HSG comes with several other downsides. Performing an HSG exposes the woman to ionizing radiation and requires the presence of a radiology department and radiographer. Additionally, HSG is usually experienced as

uncomfortable or painful with a median pain score of 5.0 (VAS, scale 0.0-10.0) (33). Discomfort or pain during HSG can be caused by various steps of the procedure; installing the instrumentation, infusion of contrast and subsequent filling of the uterine cavity with pressure build-up, and finally spillage of contrast into the peritoneal cavity.

Several analgesic interventions to support an HSG have been suggested, including topical (a spray on the cervix or intrauterine installations), oral, locally injected and intravenous pain medication. Meta-analysis only showed an effect for intravenous opioid analgesia (mean difference (MD) -3.53, 95% CI -4.29 to -2.77) and topical analgesics (MD -0.63 95% CI -1.06 to -0.19) when compared to placebo or no treatment (39). As intravenous opioids can lead to significant side effects and risks this requires continuous monitoring and is not a suitable option in most settings. The effect of topical analgesics is very small and the clinical relevance of this reduction in pain perception can be debated. Therefore, an effective method for pain relief during HSG, with a low risk profile, is needed.

A relatively novel and innovative technique for non-pharmaceutical pain relief is the use of Virtual Reality (VR). This therapy is delivered through a head-mounted device, covering the eyes and optionally the ears, to take the user into a virtual world. Viewing and interacting with the virtual world distracts the user from physical stimuli and has been proven effective in the reduction of acute pain and pain caused by medical procedures (40). Virtual Reality has the potential to improve patient experiences during HSG by providing pain relief in addition to currently used pharmacological pain relief.

Alternatives to HSG

Other visual tubal patency tests have been developed that are more tolerable, are less invasive, are easier to perform and/or are less expensive when compared to surgery or HSG.

In these tests, ionizing or echogenic contrast or coloured dye is infused into the cervix transvaginally, to detect whether there is overflow of contrast from the Fallopian tubes into the peritoneal cavity.

The most common alternatives to HSG are hysterosalpingo-foam sonography (HyFoSy), hysterosalpingo-contrast sonography (HyCoSy), transvaginal hydrolaparoscopy (THL / fertiloscopy), and magnetic resonance hysterosalpingography (MR-HSG) (41). These alternatives have been investigated and compared to HSG and/or laparoscopy with varying results and there currently is no consensus on the diagnostic quality of each of these tests. The role or place of these tests is not yet established and therefore a comparative review is warranted.

During HyFoSy, a foamy contrast liquid is inserted into the uterus transvaginally while at the same time a transvaginal ultrasound is made. The contrast contains air bubbles that give an echogenic reflection on the ultrasound monitor. As the contrast enters the genital tract, the shape of the uterine cavity and Fallopian tubes will be delineated and tubal patency can be established. This diagnostic procedure can be conducted by a gynaecologist, fertility doctor, nurse or sonographist and does not require radiation or the use of iodinated contrast

A recent large multicentre RCT, the FOAM study, compared management based on the results of HyFoSy with management based on the results of HSG (42). The study concluded that management based on the results of the HyFoSy result in an equal number of pregnancies as management based on the results of HSG (difference -1.2%; 95% CI: -3.4% to 1.5%). Due to the design of the study, the therapeutic effect of the individually used tests could not be evaluated as all participants underwent both HyFoSy and HSG in a randomized order. The mean pain score, a secondary outcome of this study, was significantly lower for HyFoSy than for HSG (3.1 standard deviation 2.2) for HyFoSy and 5.4 (standard deviation 2.5) for HSG ($p<0.001$) which is in line with previous research reporting low pain scores during HyFoSy (43).

TREATMENT OF INFERTILITY

Depending on the results of the fertility work-up, and taking into account prognostic factors for conception, a treatment plan will be determined as can be seen in Figure 1 on the right hand side. There is a large variability in treatment options and strategies between different countries and regions. While in some countries a prognostic model such as the Hunault model is used (44), other countries rely on female age or duration of infertility alone in the decision to go for expectant management or start fertility treatment. Prognostic models have been developed to guide physicians in determining the best treatment strategy for the individual couple aiming to prevent overtreatment. The prognostic model of Hunault (44) takes into account female age, duration of infertility, referral status, whether the woman is facing primary or secondary infertility and the motility of the male partners semen. The algorithm then calculates the chance of natural conception within 12 months based on the given parameters. A prognosis of ≥30% chance of natural conception within 12 months will guide physicians to counsel patients for expectant management for a period of 6-12 months (45).

PSYCHOLOGICAL IMPACT

The treatment of infertility, whether this is an advice for expectant management or medically assisted reproduction (MAR) technique, poses a high psychological burden (46). The entire path of infertility, from the wish to conceive to fertility treatments, is associated with an increased stress level and a reduced quality of life for both women and men (47, 48) (see figure 1). The uncertainty, prolonged unfulfilled wish for a child and repeated setbacks have a large impact on wellbeing in general.

Coronavirus Disease 2019 (COVID-19) led to a pandemic during the course of this PhD-track and was a big cause for stress and a reduced quality of life for people worldwide as it impacted the economy and led to environmental, health-related and political problems (49). As the path of infertility, COVID-19 came with great uncertainty and repeated setbacks. Being quarantined seems to impact women more than man; women showed a higher anxiety level and lower quality of life than man during the pandemic in a Spanish study (50).

During the pandemic, governments and health care institutions had to relocate health care workers and available resources in order to focus on the care for patients with COVID-19, and measures had to be taken to prevent further spread of the disease, which resulted in several lockdowns and a halt in healthcare for patients with non-threatening diseases. In the Netherlands, all fertility care was stopped completely in March 2020 except for patients in need of fertility preservation due to oncological indications. From May 2020, fertility care was slowly restarted but remained at a lower capacity (approximately 70% in comparison to 2019) due to the preventative measures in all hospitals and clinics up to the spring of 2022. This resulted in couples with infertility being put on hold for a prolonged period before fertility work-up could start and they found themselves in growing uncertainty. In addition, due to the lower capacity of infertility clinics, patients faced long waiting lists for fertility treatments as well. Although infertility is not a life threatening disease, it is likely that the COVID-19 pandemic has aggravated the psychological burden that infertility is associated with.

AIM OF THIS THESIS

In this thesis, we continue where the H2Oil-study (23) and it's follow-up study (51) have left us. As the fertility enhancing effect of oil-based contrast is now established for women with unexplained infertility and a low risk for tubal pathology, we want to examine the safety of this contrast, the duration of the fertility enhancing effect, and

whether the fertility enhancing effect is also present in other subgroups of infertile women. Furthermore we want to compare the diagnostic quality of the HSG with more contemporary tubal patency tests and examine an innovative method for pain reduction during HSG. As the recent COVID-19 pandemic had an profound global effect on elective health care, we wanted to investigate the psychological impact of the COVID-19 pandemic on women facing infertility.

OUTLINE

In **Chapter 2**, we start with an investigation into the duration of the fertility enhancing effect of tubal flushing with oil-based contrast in women with idiopathic infertility. With this secondary analysis of the H2Oil study and it's follow-up study, we aim to gain more insight in the mechanism of action of the fertility enhancing effect of oil-based contrast. In this chapter, we will answer the question: *Does the fertility-enhancing effect of tubal flushing during HSG change over time?*

The H2Oil study excluded three important subgroups of women: women of advanced age (39 years of age or over), women with an increased risk for tubal pathology based on their medical history and women with ovulation disorders. In **chapter 3** we present the H2Oil2 study, an ongoing international randomized controlled trial in which we study these three groups that were excluded from the H2Oil study. The research question is: *What is the effectiveness and cost-effectiveness of the use of oil-based versus water-based contrast medium during HSG in terms of live birth in women who are 39 years or older, women wo have a high risk for tubal pathology or who have an ovulation disorder?* The results of this study together with its cost-effectiveness analysis will help clinicians and policy-makers to determine what the most appropriate use of contrast is for each patient group.

In **Chapter 4** we also look into the most appropriate use of resources for specific patient groups. Using the H2Oil-study database, we tested the previously studied hypothesis that IUI has a greater beneficial effect on pregnancy chances than expectant management in couples with a low prognosis for natural conception compared to couples with a good prognosis. The research question was: *Can we replicate the finding that the benefit of IUI-ovarian stimulation compared to expectant management for couples with unexplained infertility depends on the prognosis of natural conception?*

Aside from the beneficial effect of HSG with oil-based contrast, we also look at potential downsides in this thesis. **Chapter 5** gives an overview of all literature reports on

complications after HSG with oil-based contrast and the potential differences with the safety profile of water-based contrast. It answers the question: *What are the frequencies of complications during or after an HSG with oil-based contrast in infertile women and/ or their offspring?*

Another downside of HSG is the discomfort or pain that patients experience during the procedure. In **chapter 6** we explore whether the use of virtual reality can improve the patient experience by distraction and relaxation. The research question in this chapter is: *Is Virtual Reality an effective tool to reduce procedural pain during HSG?*

After several studies investigating HSG, we will also look into other visual tubal patency tests. We are performing a review comparing the diagnostic test accuracy of the most commonly used tubal patency tests: HSG, HyFoSy, HyCoSy, THL and MR-HSG. The protocol for this study, designed to answer the question '*What is the diagnostic accuracy of the various types of visual tubal patency tests for diagnosing tubal occlusion?'* can be found in **chapter 7**.

Finally, we conduct a national survey study among women whose fertility work-up or treatment or endometriosis treatment was paused due to the COVID-19 pandemic in 2020. With this survey we want to answer the following research question: *What is the impact of the treatment pause on quality of life and quality of care of patients with infertility or endometriosis in the Netherlands?* The results are presented in **chapter 8**.

REFERENCES

1. Zegers-Hochschild F, Adamson GD, Dyer S, Racowsky C, de Mouzon J, Sokol R, et al. The International Glossary on Infertility and Fertility Care, 2017. Hum Reprod. 2017;32(9):1786-801.
2. WHO. International Classification of Diseases. 2018.
3. Mascarenhas MN, Flaxman SR, Boerma T, Vanderpoel S, Stevens GA. National, regional, and global trends in infertility prevalence since 1990: a systematic analysis of 277 health surveys. PLoS Med. 2012;9(12):e1001356.
4. Thoma ME, McLain AC, Louis JF, King RB, Trumble AC, Sundaram R, et al. Prevalence of infertility in the United States as estimated by the current duration approach and a traditional constructed approach. Fertil Steril. 2013;99(5):1324-31 e1.
5. Lancet T. Global Burden of Disease: Global Burden of Disease; [Available from: https://www.thelancet.com/gbd/about.
6. Sun H, Gong TT, Jiang, Y.T., Zhang S, Zhao YH, et al. Global, regional, and national prevalences and disability-adjusted life-yearrsr for infertility in 195 countries and territorries, 1990-2017: results froom a global burden of disease study, 2017. AGING. 2019;11:10952-91.
7. RCOG. Fertility problems: asstessment and treatment. London; 2013. Contract No.: CG 156.
8. ACOG. Infertility workup for the women's health specialist. ACOG Committee Opinion No. 781. 2019.
9. Bhattacharya S, Porter M, Amalraj E, Templeton A, Hamilton M, Lee AJ, et al. The epidemiology of infertility in the North East of Scotland. Hum Reprod. 2009;24(12):3096-107.
10. WHO. WHO Laboratory manual for the examination of human semen and sperm-cervical mucus interaction. Geneva (Switzerland); 2010.
11. Rantsi T, Land JA, Joki-Korpela P, Ouburg S, Hokynar K, Paavonen J, et al. Predictive Values of Serum Chlamydia trachomatis TroA and HtrA IgG Antibodies as Markers of Persistent Infection in the Detection of Pelvic Adhesions and Tubal Occlusion. Microorganisms. 2019;7(10).
12. Kosseim M, Brunham RC. Fallopian tube obstruction as a sequela to *Chlamydia trachomatis* infection. Eur J Clin Microbiol. 1986;5(5):584-90.
13. Luttjeboer FY, Verhoeve HR, van Dessel HJ, van der Veen F, Mol BW, Coppus SF. The value of medical history taking as risk indicator for tuboperitoneal pathology: a systematic review. BJOG. 2009;116(5):612-25.
14. Rubin IC. The nonoperative determination of patency of Fallopian tubes. JAMA. 1920;75:661-7.
15. Ross JW. Transuterine tubal insufflation (Rubin's method). J Natl Med Assoc. 1955;47:165-7.
16. Guideline-Fertility workup: Dutch Society of Obstetrics and Gynaecology; 2015 [Available from: https://richtlijnendatabase.nl/richtlijn/orienterend_fertiliteitsonderzoek_ofo/startpagina_orienterend_fertilitietsonderzoek.html.
17. Gordts S, Campo R, Rombauts L, Brosens I. Transvaginal hydrolaparoscopy as an outpatient procedure for infertility investigation. Hum Reprod. 1998;13:99-103.
18. Bergin JHE. The advantages and disadvantages of salpingography with particular reference to the use of dioddone viscous. Br J Radiol. 1951;24:93-102.
19. Bonnemain B, Guerbet M. The history of Lipiodol (1901-1994) or how a medication may evolve with the times. Rev Hist Pharm Paris. 1995;42:159-70.
20. Schoemaker J. Serial-hysterosalpingography. Amsterdam: Vrije Universiteit; 1973.
21. Perry EB. Uterosalpingography. J Natl Med Assoc. 1935;27:153-8.
22. Barnett E. The clinical value of hysterosalpingography. Clinical Radiology. 1956(3):184-96.

23. Dreyer K, van Rijswijk J, Mijatovic V, Goddijn M, Verhoeve HR, van Rooij IAJ, et al. Oil-Based or Water-Based Contrast for Hysterosalpingography in Infertile Women. N Engl J Med. 2017;376(21):2043-52.
24. Wang R, van Welie N, van Rijswijk J, Johnson NP, Norman RJ, Dreyer K, et al. Effectiveness on fertility outcome of tubal flushing with different contrast media: systematic review and network meta-analysis. Ultrasound Obstet Gynecol. 2019;54(2):172-81.
25. Wang R, Watson A, Johnson N, Cheung K, Fitzgerald C, Mol BWJ, et al. Tubal flushing for subfertility. Cochrane Database Syst Rev. 2020;10:CD003718.
26. Zhang J, Lan W, Wang Y, Chen K, Zhang G, Yang W, et al. Ethiodized poppyseed oil-based contrast medium is superior to water-based contrast medium during hysterosalpingography regarding image quality improvement and fertility enhancement: A multicentric, randomized and controlled trial. EClinicalMedicine. 2022;46:101363.
27. Roest I, Hajiyavand AM, Bongers MY, Mijatovic V, Mol BWJ, Koks CAM, et al. What is the fertility-enhancing effect of tubal flushing? A hypothesis article. J Obstet Gynaecol. 2022:1-7.
28. Johnson NP, Bhattu S, Wagner A, Blake DA, Chamley LW. Lipiodol alters murine uterine dendritic cell populations: a potential mechanism for the fertility-enhancing effect of lipiodol. Fertil Steril. 2005;83(6):1814-21.
29. Kerin JF, Williams DB, San Roman GA, Pearlstone AC, Grundfest WS, Surrey ES. Falloposcopic classification and treatment of fallopian tube lumen disease. Fertility and Sterility. 1992;57(4):731-41.
30. Johnson JV, Montoya IA, Olive DL. Ethiodol oil contrast medium inhibits macrophage phagocytosis and adherence by altering membrane electronegativity and microviscosity. Fertility and Sterility. 1992;58(3):511-7.
31. Mikulska D, Kurzawa R, Rozewicka L. Morphology of in vitro sperm phagocytosis by rat peritoneal macrophages under influence of oily contrast medium (Lipiodol). Acta Eur Fert. 1994.
32. Totorikaguena L, Olabarrieta E, Matorras R, Alonso E, Agirregoitia E, Agirregoitia N. Mu opioid receptor in the human endometrium: dynamics of its expression and localization during the menstrual cycle. Fertil Steril. 2017;107(4):1070-7 e1.
33. van Welie N, Dreyer K, van Rijswijk J, Verhoeve HR, Goddijn M, Nap AW, et al. Treatment effect of oil-based contrast is related to experienced pain at HSG: a post-hoc analysis of the randomised H2Oil study. Hum Reprod. 2019;34(12):2391-8.
34. Izumi G, Koga K, Takamura M, Bo W, Nagai M, Miyashita M, et al. Oil-Soluble Contrast Medium (OSCM) for Hysterosalpingography Modulates Dendritic Cell and Regulatory T Cell Profiles in the Peritoneal Cavity: A Possible Mechanism by Which OSCM Enhances Fertility. J Immunol. 2017;198(11):4277-84.
35. Suarez SS, Pacey AA. Sperm transport in the female reproductive tract. Hum Reprod Update. 2006;12(1):23-37.
36. Robins SA, Shapira AA. The value of hysterosalpingography: a study of 1000 cases. N Engl J Med. 1931;205:380-95.
37. Dusak A, Soydinc HE, Onder H, Ekinci F, Goruk NY, Hamidi C, et al. Venous intravasation as a complication and potential pitfall during hysterosalpingography: re-emerging study with a novel classification. J Clin Imaging Sci. 2013;3:67.
38. Roest I, van Welie N, Mijatovic V, Dreyer K, Bongers M, Koks C, et al. Complications after hysterosalpingography with oil- or water-based contrast: results of a nationwide survey. Hum Reprod Open. 2020;2020(1):hoz045.
39. Hindocha A, Beere L, O'Flynn H, Watson A, Ahmad G. Pain relief in hysterosalpingography. Cochrane Database Syst Rev. 2015(9):CD006106.
40. Mallari B, Spaeth EK, Goh H, Boyd BS. Virtual reality as an analgesic for acute and chronic pain in adults: a systematic review and meta-analysis. J Pain Res. 2019;12:2053-85.

41. Grigovich M, Kacharia VS, Bharwani N, Hemingway A, Mijatovic V, Rodgers S. Evaluating Fallopian tube patency: wha the radiologist needs to know. RadioGraphics. 2021;41(6).
42. van Welie N, van Rijswijk J, Dreyer K, van Hooff MHA, de Bruin JP, Verhoeve HR, et al. Can hysterosalpingo-foam sonography replace hysterosalpingography as first-choice tubal patency test? A randomized non-inferiority trial. Hum Reprod. 2022;37(5):969-79.
43. Dreyer K, Out R, Hompes PG, Mijatovic V. Hysterosalpingo-foam sonography, a less painful procedure for tubal patency testing during fertility workup compared with (serial) hysterosalpingography: a randomized controlled trial. Fertil Steril. 2014;102(3):821-5.
44. Hunault CC, Habbema JD, Eijkemans MJ, Collins JA, Evers JL, te Velde ER. Two new prediction rules for spontaneous pregnancy leading to live birth among subfertile couples, based on the synthesis of three previous models. Hum Reprod. 2004;19(9):2019-26.
45. Guideline-Unexplained infertility: Dutch Society of Obstetrics and Gynaecology; 2020 [Available from: https://richtlijnendatabase.nl/richtlijn/onverklaarde_subfertiliteit/startpagina_-_onverklaarde_subfertiliteit.html.
46. Cousineau TM, Domar AD. Psychological impact of infertility. Best Pract Res Clin Obstet Gynaecol. 2007;21(2):293-308.
47. Aarts JW, van Empel IW, Boivin J, Nelen WL, Kremer JA, Verhaak CM. Relationship between quality of life and distress in infertility: a validation study of the Dutch FertiQoL. Hum Reprod. 2011;26(5):1112-8.
48. Chachamovich JR, Chachamovich E, Ezer H, Fleck MP, Knauth D, Passos EP. Investigating quality of life and health-related quality of life in infertility: a systematic review. J Psychosom Obstet Gynaecol. 2010;31(2):101-10.
49. El Keshky MES, Basyouni SS, Al Sabban AM. Getting Through COVID-19: The Pandemic's Impact on the Psychology of Sustainability, Quality of Life, and the Global Economy - A Systematic Review. Front Psychol. 2020;11:585897.
50. Ferreira LN, Pereira LN, da Fe Bras M, Ilchuk K. Quality of life under the COVID-19 quarantine. Qual Life Res. 2021;30(5):1389-405.
51. van Rijswijk J, van Welie N, Dreyer K, Pham CT, Verhoeve HR, Hoek A, et al. Tubal flushing with oil-based or water-based contrast at hysterosalpingography for infertility: long-term reproductive outcomes of a randomized trial. Fertil Steril. 2020;114(1):155-62.

CHAPTER 2

How long does the fertility-enhancing effect of hysterosalpingography with oil-based contrast last?

Kimmy Rosielle‡, Nienke van Welie‡, Kim Dreyer, Joukje van Rijswijk, Cornelis B. Lambalk, Nan van Geloven, Velja Mijatovic, Ben Willem J. Mol, Rik van Eekelen, on behalf of the H2Oil Study Group

‡These authors contributed equally to this work.

Reprod Biomed Online. 2020 Dec;41(6):1038-1044

ABSTRACT

Research question

Does the fertility-enhancing effect of tubal flushing during hysterosalpingography (HSG) with oil-based contrast change over time?

Design

This was a secondary analysis of the H2Oil (long-term follow-up) study, a multicentre randomized controlled trial evaluating the effectiveness of oil-based and water-based contrast during HSG. The main outcome was ongoing pregnancy. Cox proportional hazards models for time to ongoing pregnancy were fitted over 3 years of follow-up.

Results

Data on 1107 couples were available; 550 couples had oil-based contrast and 557 water-based contrast at HSG. Ongoing pregnancy rates after 3 years were 77% and 71%, respectively. Median follow-up was 9–10 months (5th—95th percentile: <1 to 36). The hazard ratio for ongoing pregnancy for oil versus water over 3 years of follow-up was 1.26 (95% confidence interval [CI] 1.10–1.45). The scaled Schoenfeld residual plots showed a decrease in hazard ratio that was linear with log-transformed time. After including an interaction with log-transformed time, the hazard ratio immediately after HSG was 1.71 (95% CI 1.27–2.31) and reduced to no effect (hazard ratio of 1) at approximately 2 years. There was no evidence for a change in hazard ratio over time in a subgroup of women who experienced pain during HSG.

Conclusions

The hazard ratio for ongoing pregnancy of oil-based versus water-based contrast was 1.71 immediately after HSG, gradually decreasing and plateauing towards a hazard ratio of 1 (indicating no effect) after approximately 2 years. This supports the hypothesis that oil-based contrast might dislodge debris or mucus plugs from the Fallopian tubes, but this has yet to be definitively proved.

INTRODUCTION

Hysterosalpingography (HSG) is a commonly applied tubal patency test during fertility workup (1, 2). Although it was first introduced as a diagnostic test, therapeutic effects have been debated in studies for many years, especially regarding HSG with use of oil-based contrast (3).

In 2017, a multicentre randomized controlled trial (RCT) (under the name of the H2Oil study) showed that HSG using oil-based contrast resulted in a 10% higher absolute ongoing pregnancy rate within 6 months compared with the use of water-based contrast (relative risk 1.37, 95% CI 1.16–1.61) (4). Two subsequent meta-analyses confirmed these findings (5, 6). The most recent meta-analysis aimed to evaluate the long-term effects of tubal flushing; however, only three studies reported a follow-up of more than 12 months, so no definitive conclusions could be drawn (6). This emphasized the need for long-term follow-up studies.

Recently, the long-term reproductive outcomes of HSG with oil-based versus water-based contrast have been published (under the name of the H2Oil follow-up study). Over a 5-year follow-up period, HSG with oil-based contrast during fertility workup resulted in more ongoing pregnancies, more live births and a shorter time to pregnancy compared with HSG with water-based contrast (7). However, it remained uncertain whether the 5-year effect was explained by the initial effect of oil-based contrast immediately after HSG or whether the effect was long(er) lasting. Exploring the duration of this fertility-enhancing effect might provide more information on the mechanism of effect of oil-based contrast. To date, this has remained unclear.

Several potential mechanisms have been suggested. They can be categorized according to their location of action: the fallopian tube, the endometrium and the peritoneum. First, tubal flushing, i.e. mechanical flushing of debris or mucus plugs or unblocking of peritubal adhesions, can clear passage of otherwise normal fallopian tubes (8). Second, uterine bathing with oil-based contrast can enhance endometrial receptivity. Oil-based contrast is produced from poppy seed and contains opium alkaloids, which potentially interact with opioid receptors in the endometrium (9) or through alterations of the uterine immune response (10). A third potential mechanism is that oil-based contrast reduces peritoneal macrophage phagocytosis and macrophage adherence, by forming an oily layer over the macrophages changing their shape and surface configuration (11). Previous studies showed that sperm phagocytosis is inhibited *in vitro* by oil-based contrast (12, 13).

More knowledge on the duration of the fertility-enhancing effect of oil-based contrast might contribute to the understanding of the underlying mechanism. It was postulated that an effect on the endometrium or on the immune response in the peritoneum would be short lasting, and that dislodging of mucus or debris from the proximal parts of the Fallopian tubes might be painful but longer lasting (i.e. over multiple cycles). This information can contribute to the search for the mechanism underlying the fertility-enhancing effect of oil-based contrast. The present study investigated whether or not the fertility-enhancing effect of HSG using oil- versus water-based contrast would change over time.

MATERIALS AND METHODS

The H2Oil study was a multicentre RCT comparing oil-based and water-based contrast in women scheduled for HSG during their fertility workup (Netherlands Trial Register [NTR] 3270) and was approved by the Institutional Review Board of the Amsterdam University Medical Centre – Academic Medical Centre (reference 2008.362, dated 12 February 2009). The H2Oil follow-up study assessed the long-term outcomes of the H2Oil trial (NTR 6577) and was approved by the Institutional Review Board of the Amsterdam University Medical Centre – VU University Medical Centre (reference 2017.221, dated 14 June 2017).

Study details and results have previously been published (4, 7). In short, the H2Oil trial recruited a total of 1119 participants in a network of 27 hospitals in the Netherlands between 3 February 2012 and 29 October 2014 (4). Participating infertile women were aged between 18 and 39 years, had an ovulatory cycle, had a low risk of tubal pathology according to their medical history, were without known endocrine disorders and had partners that had a total motile sperm count after sperm wash of more than 3 million/ml. They had been trying to conceive for at least 1 year and were scheduled for tubal patency testing with HSG at the end of the fertility workup. After informed consent, couples were randomized for HSG with oil-based contrast or water-based contrast. In the H2Oil follow-up study, data regarding fertility treatments and pregnancies were collected until 3–5 years after randomization (7).

Study outcomes

The main outcome was ongoing pregnancy, defined as an ultrasound-confirmed positive heartbeat beyond 12 weeks of gestation. Additional to various other pregnancy outcomes, data on fertility treatments were collected. The start of follow-up was defined as 2 weeks before HSG (reflecting the first day of the menstruation before HSG). Time to pregnancy was defined as 2 weeks before HSG to the first day of menstruation before

conception leading to an ongoing pregnancy, loss to follow-up or end of study follow-up, whichever occurred first. Median follow-up was calculated as the 50th percentile in all numerical follow-up values. Pain experienced during HSG was reported using a visual analogue scale (VAS) score (range 0.0 to 10.0 in centimetres).

Primary analysis

For long-term follow-up up to 3 years, first a Cox proportional hazards model was fitted for time to pregnancy data with the randomization allocation, i.e. oil versus water, and the overall hazard. Scaled Schoenfeld residuals were then derived and plotted to visualize the change in log hazard ratio over time; a chi-squared test was applied to the residuals to test the plausibility of the proportional hazards assumption that 'the relative effect is stable over time' (14). For these tests and plots, both regular time and log-transformed time were used. The non-linear and linear interactions transformed time were added to the Cox model, and the best fitting model was determined by looking at the *P*-value for the interaction and/or whether the model had lowest Akaike information criterion (AIC) (15, 16). This best-fitting model was used to quantify the change of effect of oil- versus water-based contrast over time by estimating hazard ratios at different time points during follow-up: at 2 weeks of follow-up (which is directly after HSG), and after 1 month (which is the start of the next menstrual cycle after HSG), 3 months, 6 months, 9 months, 1 year, 2 years and 3 years.

The number of pregnancies per group per cycle during the first 6 months after HSG and their relative risks were tabulated to look for a trend over time shortly after HSG.

Sensitivity analyses

Three sensitivity analyses were conducted. For the first, it was postulated that women might experience pain at HSG when the contrast fluid removed debris or mucus plugs from their Fallopian tubes. If this were the mechanism of action, there might not be any change over time in the effect of oil-based contrast versus water-based contrast in this group. To test this, the steps from the primary statistical analysis were repeated in the subgroup of women who experienced pain during the HSG procedure, defined as a pain score of 6 points or more on the VAS.

Second, it was postulated that starting IVF at some point during follow-up might distort the effect of oil versus water given the hypothesis that oil-based contrast would flush debris or mucus plugs, as IVF bypasses the Fallopian tubes. For this reason, follow-up was censored, i.e. stopped, when couples started IVF, and the steps in the primary statistical analysis were repeated.

Third, as patient characteristics were similar between allocation groups at randomization but might differ later in follow-up, the following characteristics were adjusted for, and the following steps from the primary statistical analysis repeated: female age, duration of infertility, primary or secondary infertility, percentage of progressive motile spermatozoa, volume of semen sample, referral status by general practitioner or specialist, abnormal HSG result in terms of blockages, female smoking status and female body mass index.

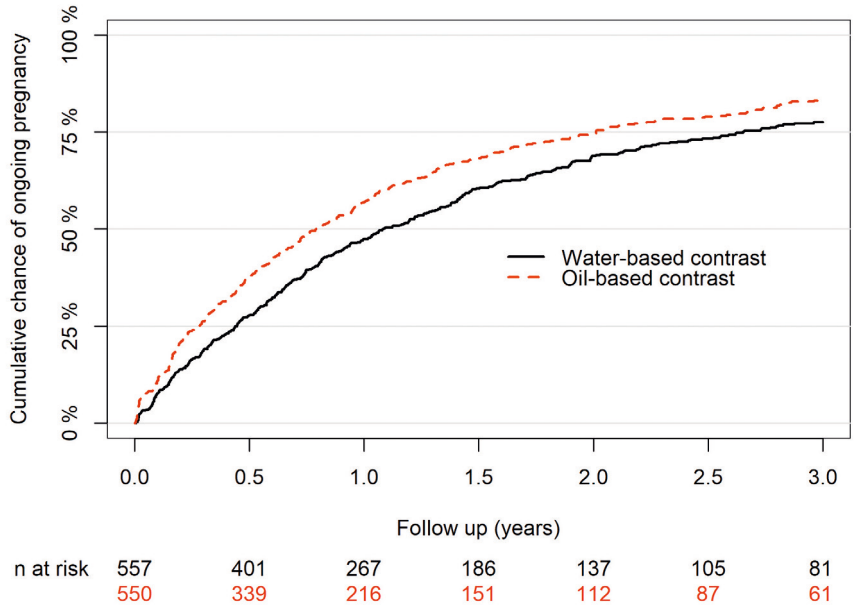
Supplementary analysis

As a supplementary analysis, the study continued with the question of whether the effect of oil contrast versus water contrast would be different for couples undergoing expectant management or receiving medically assisted reproduction (MAR), i.e. intrauterine insemination (IUI)/intrauterine insemination with ovarian stimulation (IUI-OS) or IVF/intracytoplasmic sperm injection (ICSI) as the Fallopian tubes were bypassed in IVF/ICSI. First, the follow-up data were reformatted by dividing the follow-up time into periods when couples where either pursuing expectant management, receiving IUI/IUI-OS or receiving IVF/ICSI. The start of follow-up for MAR was defined as 14 days before the first day of last menstruation previous to commencing treatment. The end of MAR follow-up was defined as the first day of the last menstruation before either ongoing pregnancy or the final insemination or embryo transfer. This aligned with the definition of time to natural conception. A treatment indicator was created that denoted which treatment (expectant, IUI/IUI-OS or IVF/ICSI) was received in which time period. Next, a Cox model was fitted with the oil versus water allocation, the MAR indicator and the interaction between these two. This model estimates the effect of oil versus water separately for expectant management, IUI/IUI-OS and IVF/ICSI.

RESULTS

The H2Oil study randomized 1119 couples. After excluding couples who conceived before receiving HSG and couples with missing follow-up or pregnancy data, data on 1107 couples were available, of which 550 couples received oil-based contrast and 557 couples received HSG with water-based contrast. Ongoing pregnancy rates at 3 years were 426 (77%) and 394 (71%), respectively. Figure 1 displays ongoing pregnancy rates as a Kaplan–Meier curve including the sample size over time. Average female age at randomization was 32.7 years (5th–95th percentile: 26.1–38.9) and median duration of infertility was 1.61 years (5th–95th percentile: 0.91–3.89). A total of 746 (67%) couples had primary infertility. The median follow-up for all couples including up to conception was 9–10 months (5th–95th percentile: <1 to 36).

Figure 1. Kaplan-Meier curve for time to an ongoing pregnancy.



Primary analysis

The hazard ratio for ongoing pregnancy estimated over 3 years follow-up of oil-based versus water-based contrast was 1.26 (95% CI 1.10–1.45). Using regular time, the scaled Schoenfeld residuals plot showed a slight decrease in hazard ratio over time from approximately 1.6 after HSG to 1.3 after 1 year (Supplementary Figure 1, test for non-proportional hazards over time: $P = 0.10$). Using log-transformed time, the decrease in log hazard ratio over time was more pronounced shortly after HSG and was statistically significant (Supplementary Figure 2, test for non-proportional hazards over time: $P = 0.02$). There thus seems evidence that the effect of oil versus water decreases over time, and that this occurs mostly within the first year after HSG.

The best fitting model included a linear term for the interaction between log-transformed time and allocation ($P = 0.02$ for interaction, an approximately 3-point decrease in AIC compared with the model with only a main treatment effect). The estimated hazard ratios at different follow-up time points are shown in Table 1. The estimated hazard ratio of oil versus water started at 1.71, gradually decreased over follow-up time and eventually plateaued around 1 after approximately 2 years of follow-up.

Table 1. Estimated HR for ongoing pregnancy of oil versus water HSG at different time points after HSG

Follow-up time	HR (95%CI) for ongoing pregnancy using log time
Immediately after HSG	1.71 (1.27-2.31)
1 month	1.57 (1.24-1.99)
3 months	1.36 (1.17-1.59)
6 months	1.25 (1.09-1.44)
9 months	1.19 (1.03-1.38)
1 year	1.15 (0.98-1.35)
2 years	1.06 (0.86-1.3)
3 years	1.00 (0.79-1.28)

CI, confidence interval; HR, hazard ratio; HSG, hysterosalpingography

The numbers of pregnancies for the first six cycles are shown in Table 2. The relative risks ranged from 1.16 to 1.83. The effect seemed to last during these six cycles but due to a small number of pregnancies per cycle, the estimates for this approach were uncertain, making it difficult to ascertain a clear trend.

Table 2. Number of pregnancies per group, per cycle

Cycle after HSG	Ongoing pregnancies oil group	Ongoing pregnancies water group	Relative Risk (95% CI)
1	47/550	31/557	1.54 (0.99-2.38)
2	39/502	30/526	1.36 (0.86-2.16)
3	41/463	24/496	1.83 (1.12-2.98)
4	24/422	23/471	1.16 (0.67-2.03)
5	22/398	19/447	1.30 (0.71-2.37)
6	23/374	19/428	1.39 (0.77-2.50)

CI, confidence interval; HSG, hysterosalpingography

Sensitivity analyses

When repeating the primary analysis only in women who were asked to judge their pain on a visual analogue scale (n = 401) and scored at least 6 points (n = 152: 73 oil and 79 water), the estimated hazard ratio over 3 years was 1.47 (95% CI 1.03–2.12). There was no evidence of a change in effect of oil versus water over time as the scaled Schoenfeld residuals plot showed a slight increase rather than a decrease, and the tests for regular and log-transformed time were not significant (Supplementary Figures 3 and 4, P = 0.88 and P = 0.71, respectively).

When censoring for IVF/ICSI, the estimated hazard ratio for oil versus water over 3 years was 1.29 (95% CI 1.11–1.50). Results from plots and tests using scaled Schoenfeld residuals were very similar to those in the primary analysis, as were the estimated hazard ratios at sequential time points (results not shown).

When adjusting for baseline characteristics, the estimated hazard ratio for oil versus water over 3 years was 1.30 (95% CI 1.13–1.50). Results from plots and tests using scaled Schoenfeld residuals were very similar to those in the primary analysis, as were the estimated hazard ratios at sequential time points (results not shown).

Supplementary analysis

There was found no evidence that the effect of oil versus water was different for expectant management, IUI/IUI-OS or IVF/ICSI: the interaction between oil versus water allocation and MAR treatment was not significant ($P = 0.39$) and did not lead to a better fit in terms of AIC.

DISCUSSION

Evidence was found that the hazard ratio for ongoing pregnancy after an HSG with oil-based contrast versus water-based contrast was highest shortly after HSG and then gradually decreased. This change was best described as linear with log-time, decreasing from a hazard ratio of 1.71 to 1, i.e. no effect, after approximately 2 years. In the subgroup of women who experienced pain during HSG, which might be because flushing dislodged debris or mucus plugs in their Fallopian tubes, there was found no evidence for a change in hazard ratio over time.

A decreasing hazard ratio over time could be due to three potential mechanisms. The first is that, for each woman, the hazard ratio for the effect of oil-based versus water-based contrast diminishes over time. Second, a decreasing hazard ratio may also be explained by heterogeneity of treatment effect, meaning that the oil contrast may have a beneficial effect that is stable over time in only a subgroup of women. These women quickly conceived and, due to selection, at later time points the treatment effect was only evaluated in women for whom it was not beneficial, leading to a lower hazard ratio. A third possible explanation is unobserved heterogeneity, meaning that pregnancy chances varied between couples due to factors unknown to the authors. Even with a treatment effect that is constant over time and similar for all women, unobserved heterogeneity may lead to attenuation of the hazard ratio towards 1 over time (17, 18). When adjusting for baseline characteristics that are known prognostic factors, the results did not differ

from the primary analysis. In addition, it can be argued that, in the first year after HSG, unobserved heterogeneity might not yet play a role and that the observed decrease in hazard ratio can be explained by either one of the former mechanisms.

In terms of strengths and limitations, this secondary analysis was performed using data from a well-designed multicentre RCT with a long follow-up period of 3–5 years. Using an objective outcome measure, ongoing pregnancy, the risk of bias was minimal. Only women with unexplained or mild male infertility were included; they were below 39 years of age, did not have known endocrinological disorders and had a low risk of tubal pathology based on their medical history. Therefore, it is questionable whether the findings are generalizable to infertile women who do not share these features. Additionally, it should be noted that the main outcome in this study was ongoing pregnancy, whereas in clinical practice live birth is the desired outcome. However, there are several reasons to justify the use of ongoing pregnancy as a proxy for live birth in fertility research (19).

The finding that the fertility-enhancing effect of oil-based contrast lasts for a substantial amount of time promotes the hypothesis that the mechanism of action lies in the Fallopian tubes, implying that tubal flushing during HSG dislodges debris, mucus plugs or small adhesions in the proximal parts of the Fallopian tubes, thereby resolving an 'unexplained' fertility factor (8, 20). Our findings are less consistent with the other suggested mechanisms, as it was postulated here that an effect in the endometrium or alteration in the immune response in the peritoneum would be a temporary effect.

Making the assumption that the HSG using oil-based contrast does dislodge debris, mucus, etc. from the Fallopian tubes, this would mean that they are essentially 'cured', which here means that their tubes are once more fully operational and they are back to their 'normal' state of fertility. However, the 'normal' fertility potential varies considerably between women (21, 22). This inherent difference between women in terms of their chance of conception might explain why not all women in this subgroup conceive within the first couple of cycles after being 'cured': some of them with lower potentials take much longer, for example more than 1 year, to conceive.

Additionally, it has been postulated that oil could emulsify debris in Fallopian tubes, facilitating the removal of debris more efficiently (23). Furthermore, the two contrast media have many differences in chemical and physical characteristics, for example oil-based contrast (Lipiodol Ultra-Fluid®, Guerbet, France) has a lower viscosity. The oil-based contrast also contains a higher iodine concentration than the water-based contrast used here (Telebrix Hystero®, Guerbet, France). There is currently very limited evidence regarding the impact of these differences, and future studies are needed.

Although a previous analysis did not identify characteristics that were associated with a greater or lesser effect of oil-based contrast compared with water-based contrast, the hypothesis regarding the Fallopian tubes is supported by a recent analysis of perceived pain during HSG (20, 24). Women who reported a VAS score of 6.0 or more were found to benefit more from oil-based contrast (49.4% versus 29.6%, Relative Risk 1.7, 95% CI 1.1–2.5) (20). In the current study, there was no evidence that the effect of oil versus water contrast decreased over time for women who had a VAS score of 6 or higher. Thus, in addition to the effect of oil contrast being dependent on VAS score, this provides some evidence to support the theory of dislodging debris or mucus plugs in the Fallopian tubes, as that effect is likely to be (semi-)permanent. However, the small sample sizes for this sensitivity analysis must be acknowledged as the VAS score was not measured in all patients (*n* = 401) and only 152 of those scored 6 or higher. The previously mentioned second mechanism might also explain the decrease in hazard ratio that was found in the whole cohort: that there is only an effect in the group in which debris was dislodged and pain was felt, and as their chances increase, they conceive and drop out of the cohort.

This secondary analysis was performed to understand the biological mechanism underlying the fertility-enhancing effect of tubal flushing and to evaluate how long it is beneficial for infertile women. However, the study emphasized the complexity of 'unexplained infertility' with multiple unknown aspects. All studied women were below 39 years of age, had a regular ovulatory cycle and had a low perceived risk of tubal pathology, so it is unclear what causes infertility in these women. Women who did not conceive within 2 years after HSG no longer benefited from the oil-based contrast. This may support the hypothesis that tubal flushing using oil-based contrast dislodges debris or mucus plugs from the proximal parts of the tubes and that, after 2 years, most of these women conceived. This hypothesis can be further explored by studies in which the pressure build-up of oil-based and water-based contrast during HSG, resulting in dislodgment of material such as debris and mucus plugs, is investigated. Furthermore, future studies are needed to evaluate whether HSG with oil-based contrast before IVF has an fertility-enhancing effect, and to assess whether the fertility-enhancing effect of oil-based contrast is also present in women above 39 years of age with a diminished ovarian reserve, women who have ovulation disorders or women at high risk of tubal pathology.

CONCLUSION

The fertility-enhancing effect of oil-based contrast versus water-based contrast after HSG in terms of the hazard ratio was highest shortly after HSG and then decreased,

although the effect still seemed to be present for at least 1 year after tubal flushing. After approximately 2 years there was no beneficial effect. Additionally, in women who experienced pain during HSG, the effect might last longer. The current results favour the hypothesis that oil-based contrast might dislodge debris or mucus plugs from the Fallopian tubes, and contradicts other locations of action, i.e. the endometrium or the peritoneum, although this has yet to be proven definitively. The findings can be used to further investigate unexplained infertility and to counsel couples with unexplained infertility that they might still conceive naturally after HSG and that treatment could be delayed for a period of time.

REFERENCES

1. ACOG. Infertility workup for the women's health specialist. ACOG Committee Opinion No. 781. 2019.
2. RCOG. Fertility problems: asstessment and treatment. London; 2013. Contract No.: CG 156.
3. Mohiyiddeen L, Hardiman A, Fitzgerald C, Hughes E, Mol BW, Johnson N, et al. Tubal flushing for subfertility. Cochrane Database Syst Rev. 2015(5):CD003718.
4. Dreyer K, van Rijswijk J, Mijatovic V, Goddijn M, Verhoeve HR, van Rooij IAJ, et al. Oil-Based or Water-Based Contrast for Hysterosalpingography in Infertile Women. N Engl J Med. 2017;376(21):2043-52.
5. Fang F, Bai Y, Zhang Y, Faramand A. Oil-based versus water-based contrast for hysterosalpingography in infertile women: a systematic review and meta-analysis of randomized controlled trials. Fertil Steril. 2018;110(1):153-60.
6. Wang R, van Welie N, van Rijswijk J, Johnson NP, Norman RJ, Dreyer K, et al. Effectiveness on fertility outcome of tubal flushing with different contrast media: systematic review and network meta-analysis. Ultrasound Obstet Gynecol. 2019;54(2):172-81.
7. van Rijswijk J, van Welie N, Dreyer K, Pham CT, Verhoeve HR, Hoek A, et al. Tubal flushing with oil-based or water-based contrast at hysterosalpingography for infertility: long-term reproductive outcomes of a randomized trial. Fertil Steril. 2020.
8. Kerin JF, Williams DB, San Roman GA, Pearlstone AC, Grundfest WS, Surrey ES. Falloposcopic classification and treatment of fallopian tube lumen disease. Fertility and Sterility. 1992;57(4):731-41.
9. Totorikaguena L, Olabarrieta E, Matorras R, Alonso E, Agirregoitia E, Agirregoitia N. Mu opioid receptor in the human endometrium: dynamics of its expression and localization during the menstrual cycle. Fertil Steril. 2017;107(4):1070-7 e1.
10. Johnson NP, Bhattu S, Wagner A, Blake DA, Chamley LW. Lipiodol alters murine uterine dendritic cell populations: a potential mechanism for the fertility-enhancing effect of lipiodol. Fertil Steril. 2005;83(6):1814-21.
11. Johnson JV, Montoya IA, Olive DL. Ethiodol oil contrast medium inhibits macrophage phagocytosis and adherence by altering membrane electronegativity and microviscosity. Fertility and Sterility. 1992;58(3):511-7.
12. Boyer P, Territo MC, de Ziegler D, Meldrum DR. Ethiodol**Savage Laboratories, Melville, New York. inhibits phagocytosis by pelvic peritoneal macrophages. Fertility and Sterility. 1986;46(4):715-7.
13. Mikulska D, Kurzawa R, Rozewicka L. Morphology of in vitro sperm phagocytosis by rat peritoneal macrophages under influence of oily contrast medium (Lipiodol). Acta Europea Fertilitatis. 1994;25(3):203-6.
14. Grambsch PM, Therneau TM. Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. Biometrika. 1994;81(3):515-26.
15. Akaike H. A new look at the statistical model identification. IEEE transactions on automatic control. 1974;AC-19(6):716-23.
16. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in medicine. 1996;15(4):361-87.
17. Hernan MA. The hazards of hazard ratios. Epidemiology. 2010;21(1):13-5.
18. Aalen OO, Cook RJ, Roysland K. Does Cox analysis of a randomized survival study yield a causal treatment effect? Lifetime data analysis. 2015;21(4):579-93.

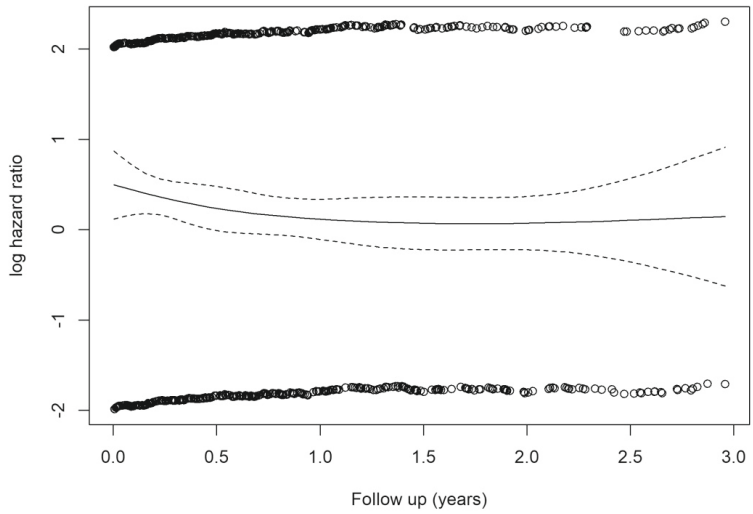
19. Braakhekke M, Kamphuis EI, Dancet EA, Mol F, van der Veen F, Mol BW. Ongoing pregnancy qualifies best as the primary outcome measure of choice in trials in reproductive medicine: an opinion paper. Fertil Steril. 2014;101(5):1203-4.
20. van Welie N, Dreyer K, van Rijswijk J, Verhoeve HR, Goddijn M, Nap AW, et al. Treatment effect of oil-based contrast is related to experienced pain at HSG: a post-hoc analysis of the randomised H2Oil study. Hum Reprod. 2019;34(12):2391-8.
21. te Velde ER, Pearson PL. The variability of female reproductive ageing. Hum Reprod Update. 2002;8(2):141-54.
22. van Eekelen R, Scholten I, Tjon-Kon-Fat RI, van der Steeg JW, Steures P, Hompes P, et al. Natural conception: repeated predictions over time. Hum Reprod. 2017;32(2):346-53.
23. Watson A, Vandekerckhove P, Lilford R, Vail A, Brosens I, Hughes E. A meta-analysis of the therapeutic role of oil soluble contrast media at hysterosalpingography: a surprising result? Fertility and Sterility. 1994;61(3):470-7.
24. van Rijswijk J, van Welie N, Dreyer K, Tajik P, Lambalk CB, Hompes P, et al. Tubal flushing with oil- or water-based contrast medium: can we identify markers that indicate treatment benefit? Hum Reprod Open. 2019;2019(3):hoz015.

Appendix: details H2Oil Study Group members

Petra Bourdrez**,** Jan Peter de Bruin, Angelique J.C.M. van Dongen, Annette E.J. Duijn, Anna P. Gijsen, Mariëtte Goddijn, Ron J.T. van Golde, Cathelijne F. van Heteren, Annemieke Hoek, Machiel H.A. van Hooff, Angelo B. Hooker, Mesrure Kaplan, Cornelia H. de Koning, Marieke J. Lambers, Alexander Mozes, Annemiek W. Nap, Marie J. Pelinck, Henrike G.M. Rijnsaardt-Lukassen, Ilse A.J. van Rooij, Alexander V. Sluijmer, Jesper M. J. Smeenk, Catharina C.M. Timmerman, Maaike A.F. Traas, Rachel Tros, Gijsbertus A. van Unnik, Harold R. Verhoeve

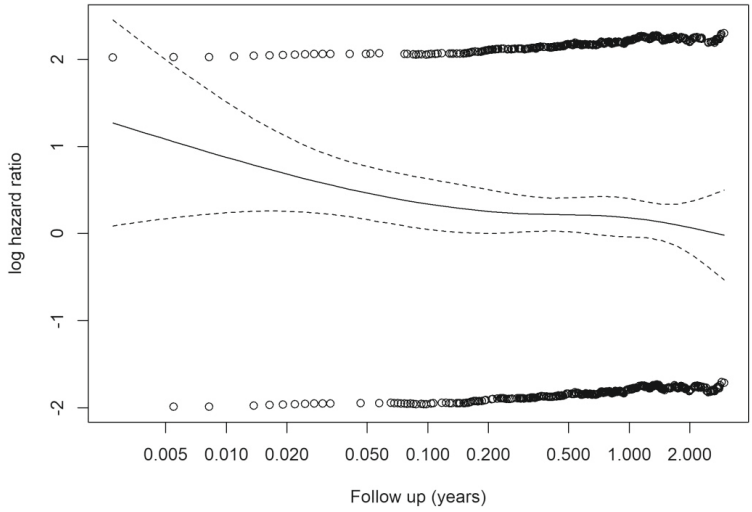
SUPPLEMENTARY MATERIALS

Supplementary Figure 1. Scaled Schoenfeld residuals plot for the basic Cox model using regular time.



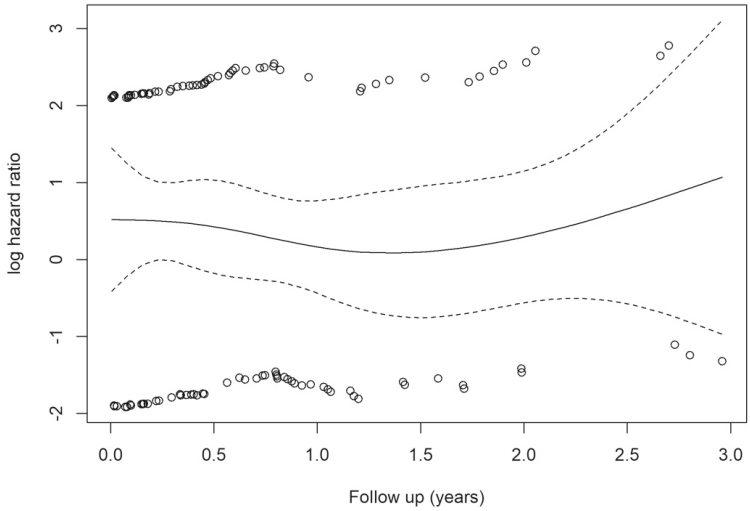
Log hazard ratio for ongoing pregnancy estimated over three years follow-up of oil-based versus water-based contrast shown by Scaled Schoenfeld residuals plot for the basic Cox model using regular time. Dotted lines represent the 95% confidence interval. Test for non-proportional hazards over time: p=0.10.

Supplementary Figure 2. Scaled Schoenfeld residuals plot for the basic Cox model using log-transformed time.



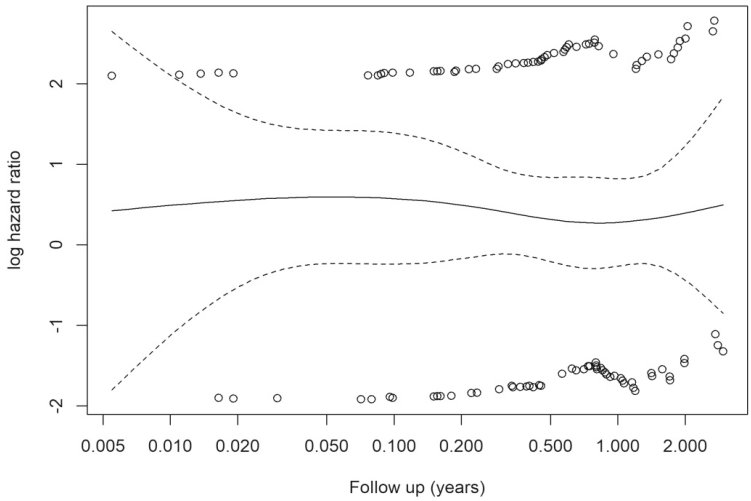
Log hazard ratio for ongoing pregnancy estimated over three years follow-up of oil-based versus water-based contrast shown by Scaled Schoenfeld residuals plot for the basic Cox model using log-transformed time. Dotted lines represent the 95% confidence interval. Test for non-proportional hazards over time: p=0.02.

Supplementary Figure 3. Scaled Schoenfeld residuals plot for the basic Cox model using regular time in a subgroup of women who experienced pain (visual analogue scale of 6 points or more) during HSG.



Log hazard ratio for ongoing pregnancy estimated over three years follow-up of oil-based versus water-based contrast shown by Scaled Schoenfeld residuals plot for the basic Cox model using regular time in a subgroup of women who experienced pain during HSG. Dotted lines represent the 95% confidence interval. Test for non-proportional hazards over time: p=0.88.

Supplementary Figure 4. Scaled Schoenfeld residuals plot for the basic Cox model using log-transformed time in a subgroup of women who experienced pain (visual analogue scale of 6 points or more) during HSG.



Log hazard ratio for ongoing pregnancy estimated over three years follow-up of oil-based versus water-based contrast shown by Scaled Schoenfeld residuals plot for the basic Cox model using log-transformed time in a subgroup of women who experienced pain during HSG. Dotted lines represent the 95% confidence interval. Test for non-proportional hazards over time: p=0.71.

CHAPTER 3

Oil-based versus water-based contrast media for hysterosalpingography in infertile women of advanced age, with ovulation disorders or a high risk for tubal pathology: study protocol of a randomized controlled trial (H2Oil2 study)

Kimmy Rosielle, Danah Kamphuis, Nienke van Welie, Inez Roest, Alexander Mozes, Evert J. P. van Santbrink, Tine van de Laar, Angelo B. Hooker, Dana A. G. Huppelschoten, Wentao Li, Marlies Y. Bongers, Jaap Stoker, Madelon van Wely, Carolien Koks, Nils C. B. Lambalk, Anne Hemingway, Ben W. J. Mol, Kim Dreyer and Velja Mijatovic

BMC Women's Health. 2022 Apr 22(1):123

ABSTRACT

Background

In women with unexplained infertility, tubal flushing with oil-based contrast during hysterosalpingography (HSG) increases ongoing pregnancy and subsequent live birth rates when compared to tubal flushing with water-based contrast. It is currently unclear whether an HSG with oil-based contrast also results in more ongoing pregnancies and live births in women of advanced age, women with ovulation disorders, and women with potential tubal pathology when compared to an HSG with water-based contrast.

Methods

We plan an international, multicentre, open-label, randomized controlled trial (RCT) studying three groups of infertile women who have an indication for tubal patency testing according to their treating physician and additionally; (1) are 39 years of age or older, (2) have an ovulation disorder or (3) have a high risk for tubal pathology based on their medical history. Women with an allergy for iodinated contrast medium are excluded, as are women with diabetes, hyperprolactinemia or untreated hyper- or hypothyroidism, and women with a partner with severe male infertility. After informed consent, women will be randomly allocated to the intervention, tubal flushing with the use of oil-based contrast during HSG or the control group, tubal flushing with the use of water-based contrast during HSG in a 1:1 ratio by the web-based system Castor. The primary endpoint will be ongoing pregnancy leading to live birth with conception within six months after randomization. Secondary outcomes are other pregnancy outcomes, used fertility treatments, adverse events and cost-effectiveness. Based on the expected ongoing pregnancy rate of 17% in the control group and 27% in the intervention group, the sample size will be 930 women (465 per group). Study inclusion is expected to be complete in four years.

Discussion

This multicentre RCT will establish whether, for women of advanced age, women with ovulatory disease, and women who have a high risk for tubal pathology, there is a fertility enhancing effect of tubal flushing with oilbased contrast during HSG and whether the use of this contrast medium is cost-effective.

Trial Registration

The study was prospectively registered in the Netherlands Trial Register on August 1st 2019 as ‘H2Oil2’ (reference number NL7925, https:// www. trial regis ter. nl/ trial/ 7925).

BACKGROUND

Infertility, defined as the inability to conceive within 12 months of unprotected intercourse, was estimated to effect approximately 48.5 million couples worldwide in 2010 (1, 2). Fertility work-up includes the medical investigation into the various causes of male and female infertility that have been identified. With male factor infertility referring to the (relative) absence of functioning spermatozoa, female infertility can refer to ovulation disorders, diminished ovarian reserve, tubal factor infertility, and uterine factor infertility (3). Ovulation disorders can have various causes, of which polycystic ovary syndrome (PCOS) is the most prevalent (4). Diminished ovarian reserve plays an increasing role in infertility as couples have been postponing their wish to conceive over recent decades (5). Tubal factor infertility can be caused by current or past pelvic inflammatory disease such as an infection with *Chlamydia trachomatis*, previous pelvic surgery, peritonitis or endometriosis (6, 7). Uterine factor infertility can consist of anatomical anomalies, such as congenital uterine anomalies, or intrauterine pathology such as polyps, myomas or adhesions (8). In up to 30% of couples, the fertility work-up shows no abnormalities and this is classified as unexplained infertility (9, 10).

Part of the fertility work-up is assessing the risk for tubal pathology and if indicated, a tubal patency test can be performed. Hysterosalpingography (HSG) is traditionally used as first choice tubal patency test to rule out tubal pathology (3, 4). Although HSG was initially introduced as a diagnostic test, therapeutic effects of tubal flushing, especially with oil-based contrast, have been studied extensively (11-14). The most recent review (14) included six studies comparing HSG with oil-based contrast and water-based contrast. Three of these studies reported on live birth but meta-analysis could not be performed due to heterogeneity of the population.

The largest study among these randomized controlled trials is the H2Oil study (15). This study was conducted to investigate the difference in ongoing pregnancy rates between tubal flushing during HSG with the use of oilbasedand water-based contrast, in couples with unexplained or mild male infertility. This study excluded women aged 39 years or older, women with ovulation disorders, and women who had a high risk for tubal pathology. The H2Oil study showed a significant increase in ongoing pregnancies as well as live births within 6 months after HSG with oil-based contrast when compared to HSG with water-based contrast (relative risk (RR) 1.37, 95% confidence interval (CI) 1.16–1.61; $P < 0.001$ for ongoing pregnancy and RR 1.38, 95% CI 1.17 to 1.64; $P < 0.001$ for live birth) (15). The long term follow-up of this study demonstrated that the fertility enhancing effect of oil-based contrast is still present five years after HSG (cumulative ongoing pregnancy rates 80.0% after oil-based contrast, 75.0% after water-based contrast, RR

1.07, 95% CI 1.00 to 1.14, cumulative live birth rates 74.8% after oil-based contrast, 67.3% after water-based contrast, RR 1.11; 95% CI 1.03 to 1.20) (16). This study additionally demonstrated that HSG with oil-based contrast leads to a significantly shorter time-to-pregnancy compared to HSG with water-based contrast (10.0 vs 13.7 months; hazard ratio 1.25; 95% CI 10.9 to 1.43). Robust studies investigating the fertility enhancing effect of oil-based contrast during HSG in women who were 39 years of age or older, women with ovulation disorders, and women who have a high risk for tubal pathology are lacking.

In the current socio-economic climate, where health care costs are increasing and the importance of evidence-based health care is underlined, the results of previously mentioned studies among couples with unexplained infertility cannot be extrapolated to couples with other types of infertility and therefore separate evaluation is needed.

This randomized controlled trial aims to investigate the effectiveness and cost-effectiveness of the use of oil-based versus water-based contrast during HSG, in women with previously unevaluated causes of infertility: women who are 39 years of age or older, women with ovulation disorders, and women who have a high risk for tubal pathology.

METHODS

This international, multicentre, randomized controlled trial will be performed in university, teaching and non-teaching hospitals in the Netherlands and the United Kingdom. A list of currently participating hospitals is available as Supplementary file 2. The trial has obtained ethical approval by the Institutional Review Board (IRB) of the Amsterdam UMC location Vrije Universiteit (registration number 2018.289), the Research Ethics Committee London Harrow (20/LO/0607), and the board of directors of all participating centres.

Participants

Women who are scheduled for an HSG as part of their fertility work-up can participate if they meet at least one of the following criteria: (1) women who are 39 years of age or older, (2) women who have an ovulation disorder (ovulation disorders will be defined as less than eight menstrual cycles per year), or (3) have a high risk for tubal pathology (high risk will be defined as a past chlamydia infection, pelvic inflammatory disease, peritonitis, known endometriosis and/or pelvic surgery including tubectomy for ectopic pregnancy). In order to ensure adequate sample sizes in all three groups of participants, women meeting more than one criterion will be included according to the criterion that comes first in ranking. The ranking is based on the expected prevalence of the three subgroups of women within the study population, with the lowest expected prevalence

highest in ranking. Women will be excluded if they have an endocrine disorder known to decrease natural pregnancy chances (e.g. diabetes, unregulated hypothyroidism or hyperthyroidism), iodine contrast medium allergy or a male partner with severe infertility (a pre-washed total motile sperm count below three million sperm per millilitre).

Randomization and blinding

Infertile couples will be screened in the outpatient clinic by their attending physician. Eligible women will be informed by a dedicated research nurse or physician in their centre. Women who give written consent will be randomized for HSG with oil-based contrast (intervention group) or with water-based contrast (control group) in a 1:1 ratio, using a permuted block design with block sizes varying from 4–8 cases. Randomization will be performed using the web-based program Castor EDC (Castor Electronic Data Capture, Ciwit BV, Amsterdam, the Netherlands), and stratified according to centres and by infertility diagnosis (age 39 years or older, ovulation disorder, and high risk for tubal pathology). The trial is not blinded with respect to participants and health care professionals since the allocation will be evident in further fertility workup. Oil-based contrast has a higher iodine concentration than water-based contrast and together with its hydrophobic qualities this makes oil-based contrast easily distinguishable from water-based contrast on X-ray or fluoroscopy images (17). The primary outcome is objective, and therefore we do not expect that lack of blinding will influence the findings.

Intervention

The HSG procedure will be performed after cessation of menstrual bleeding or after progesterone-induced vaginal bleeding in case of anovulation. With use of a cervical vacuum cup, a metal cannula (hysterophore), an acorn cannula or an HSG balloon catheter, the iodinated oil- or water-based contrast medium will be infused into the uterine cavity and several radiographs will be taken to visualize the uterine cavity and Fallopian tubes according to local protocols. The procedure will be discontinued if signs of intravasation are visible on the radiographs, as intravasation of specifically oil-based contrast might lead to oil-embolisms, a known and potentially dangerous complication (18). The results of the HSG will describe whether the Fallopian tubes are patent and whether there are any visual abnormalities in the cervix, uterine cavity, Fallopian tubes or peritoneal cavity. In the intervention arm, the HSG will be performed with a maximum of 15 millilitre of oil-based contrast (Lipiodol Ultra Fluid®, Guerbet, Villepinte, France) to minimize the chance of (temporary) hypo- or hyperthyroidism (18). In the comparator arm, the HSG will be performed with water-based contrast medium (iodixanol, Visipaque®, General Electric Healthcare, Buc, France), for which no maximum dosage is advised. The batch number and expiration date of the used flasks of contrast medium will be reported for the purpose of drug accountability.

Pre- and post-HSG use of analgesics and antibiotics, and subsequent management will be performed according to local protocol. Women will receive the adjusted Amsterdam Preoperative Anxiety and Information Scale (APAIS) questionnaire prior to their HSG procedure to score their anxiety score prior to HSG and to be able to relate this to their pain level during HSG (19). Immediately after the procedure they will be asked to score their pain using a Visual Analogous Scale (VAS) ruler (ranging from 0.0 to 10.0 cm).

The choice of fertility treatments will be based on the results of the fertility work-up (including the outcome of the HSG) according to the Dutch fertility Guideline or the Clinical guideline by the National Institute for Healthcare and Excellence (NICE) (4, 20-22). Anovulatory women will be offered ovulation induction, and women aged 39 years or older may be offered Intra Uterine Insemination (IUI) or In Vitro Fertilization (IVF). In case of suspected uni- or bilateral tubal pathology, women can be scheduled for IVF or a diagnostic or therapeutic laparoscopy followed by IVF if bilateral tubal occlusion is confirmed, according to the local protocol of the participating centres. Women with a high risk for tubal pathology, but without tubal pathology at HSG or laparoscopy, and with a regular menstrual cycle who are below 39 years of age will be advised expectant management or IUI, guided by their calculated prognosis for natural conception using the model of Hunault or other local protocols (23). For women aged 39 or over, the Hunault prognostic model is not verified and these women will often be advised IUI or IVF immediately.

As the compared strategies (HSG with oil-based contrast versus HSG with water-based contrast) are already applied in current practice, no additional risks or burdens are expected for participating women.

Outcomes

The primary outcome of this trial is ongoing pregnancy leading to live birth, with conception within six months after randomization. Ongoing pregnancy will be defined as an intrauterine pregnancy with heartbeat on ultrasound examination at twelve weeks of gestation, live birth as a live born neonate beyond 24 weeks of gestation. Secondary outcomes will include clinical pregnancy (ultrasound confirmed intrauterine gestational sac), ongoing pregnancy, miscarriage (loss of clinical or ongoing pregnancy or diagnosis of a pregnancy without positive foetal heartbeat before twelve weeks gestation), ectopic pregnancy (ultrasound or surgically confirmed extra-uterine pregnancy). Pregnancy complications, complications of HSG such as intravasation, infection and hypo- or hyperthyroidism, and a cost-effectiveness analysis will also be part of the secondary outcomes. Our hypothesis is that HSG with oil-based contrast will increase ongoing pregnancy rates and will reduce time to ongoing pregnancy in all three subgroups,

thereby reducing the need for assisted reproductive technology (ART) and thus lowering the costs. In addition, we will study the procedural discomfort or pain and relate this to pre-procedural anxiety, using a modified APAIS questionnaire (19).

Follow-up

Data on fertility treatments and pregnancy outcomes will be collected until six months after randomization in a structured electronic case report form using Castor EDC. If a pregnancy occurs within six months, the outcome of the pregnancy will be followed.

If the necessary information cannot be extracted from the medical record, women will receive a digital questionnaire about treatment and pregnancy outcomes or they will be contacted by a dedicated researcher to conduct the follow-up questionnaire by phone. All participating women will receive a digital questionnaire on productivity-loss and health care costs (iPCQ) six months after randomization (24) (see Table 1).

Table 1. SPIRIT figure. t–1: Prior to inclusion; t0: Study inclusion; t1: HSG procedure, usually within 4 weeks of inclusion (t0); t2: end of initial follow-up 6 months after randomization; t3: pregnancy follow-up, at last 9 months after t2.

	Study Period				
	Enrolment	**Allocation**	**Post-allocation**		
Time-point	*t-1*	*t0*	*t1*	*t2*	*t3*
ENROLMENT					
Eligibility screen	X				
Informed consent	X				
Allocation		X			
INTERVENTIONS					
HSG with oil-based contrast			X		
HSG with water-based contrast			X		
ASSESSMENTS					
Demographics		X			
APAIS			X		
Pain score (VAS)			X		
HSG procedure and results			X		
Adverse effects			X		
iPCQ				X	
Treatments				X	
Pregnancies				X	
Pregnancy follow-up					X

Sample size

Our hypothesis is that HSG with oil-based contrast will increase the live birth rate with 10% in three infertility groups: (1) women aged 39 years or older, (2) women with ovulation disorders and (3) women who have a high risk for tubal pathology. To detect an increase of 10% in live birth rate from 17 to 27%, 395 women per group are needed (alpha 1%, beta 20%, Z-test with unpooled variances as calculated in PASS 2020). Anticipating a loss to follow-up rate of 15%, the total number of participants required is 930 (465 in each arm of the trial). With this number we have 80% power to study the 10% difference in live birth rate in the intervention versus control group in the stratified design using Cochrane-Mantel-Haenzel.

Statistical analysis

Categorical data will be reported as absolute numbers and percentages. Normally distributed continuous variables will be summarized as means with standard deviations, and non-normally distributed continuous variables will be reported as medians with interquartile ranges. The primary analyses will done according to the intention to treat (ITT) principle, including all randomised women. Differences in live births will be expressed as crude and stratification adjusted risk ratio and absolute risk difference with associated 95% and 99% confidence intervals (CI) using log-linear binomial regression. We will construct Kaplan-Meier curves, estimating the cumulative probability of conception leading to live birth over time and use the log-rank test to assess differences. Additionally, we will do a cox proportional hazard analysis to evaluate the difference in primary outcomes over time while accounting for the subgroups and evaluating presence of interaction. Continuous outcomes will be measured at multiple time-points and will be analysed with the use of linear mixed models. We will subsequently compare intervention and control treatment within the stratified groups (1) women aged 39 years or older, (2) women with ovulation disorders and (3) women who have a high risk for tubal pathology. Women meeting more than one criterion will be included in the group that comes first in ranking as described earlier. Within these stratified groups pregnancy outcomes will be expressed as risk ratio, risk difference and hazard ratio with 95% CI.

Cost-effectiveness analysis

The average costs and effects of tubal flushing during HSG with oil-based contrast and water-based contrast during fertility work-up will be compared. Total costs of the HSG, fertility treatments and fertility outcomes (collected using the eCRF) will be evaluated after a follow-up of six months after randomization. These data will be used to calculate the direct medical costs.

Societal costs will be measured using digital iPCQ questionnaires after six months of follow-up (24). Cost categories that will be included are: (1) healthcare costs (primary and secondary care, complementary care and home care); (2) lost productivity costs (absenteeism from paid and unpaid work, and presentism) and (3) patient costs (informal care and other care services paid for by patients themselves).

Valuation for participants from the Netherlands and the United Kingdom will be according to their respective national guidelines (25, 26). For the valuation of health care utilization, lost productivity and informal care, standard costs for the Netherlands and the United Kingdom will be used. Medication use will be valued using prices of the Royal Dutch Society for Pharmacy (www.KNMP.nl) and the NICE British National Formulary (www.bnf.nice.org.uk). Patient and family costs other than informal care will be valued using self-reported prices. For the valuation of absenteeism from paid work, the friction cost approach will be used.

Safety monitoring

The IRB determined that the study related risk for participants is very low. A Data Safety Monitoring Committee was therefore not deemed necessary. An interim analysis is not planned. All adverse events (AEs) occurring within one month after HSG will be reported to the IRB by line listing yearly. Additionally, adverse neonatal outcomes such as a congenital anomaly or birth defect will be reported as severe adverse events (SAEs). SAEsoccurring within one month after HSG will be reported to the IRB immediately through the Dutch national web portal ToetsingOnline. SAEs occurring in participants outside of the Netherlands will additionally be processed according to local regulations. All SAEs will be followed until they have abated, until a stable situation has been reached or the patient was discharged.

Data management and monitoring

Patient information will be filled in anonymously based on randomization number. Linking personal data to the study number can only be performed in the local participating centres. Written informed consent forms are stored in the local participating centre, all forms and data will be archived for 25 years in the participating centres according to GCP and local regulations. Monitoring of study processes will be done according to national and international guidelines by an independent study monitor (27). Annual safety reports will be sent to the accredited IRB and the competent authority.

See Supplementary Table 1 for the completed WHO Trial Registration Data Set (28).

DISCUSSION

Increasing female age is one of the main causes of infertility in the twenty-first century, with >50% of the women undergoing IVF being over 35 years of age (29, 30). Anovulation and tubal pathology are also important causes of infertility. As a consequence, the results of this H2Oil2 study are estimated to be applicable to more than 50% of infertile women seen in fertility clinics.

This multicentre randomized controlled trial will generate insight in the potential fertility enhancing effect of tubal flushing using oil-based contrast during HSG in infertile women who are 39 years or older, women with ovulatory disease, and women with a high risk for tubal pathology. The generated evidence can guide clinicians and policy makers to decide which subgroups of women will benefit from an HSG with oil-based contrast as a therapeutic intervention, whether the intervention is cost-effective and if the risk of adverse events is acceptable.

Strengths and limitations

The proposed study is specifically designed to detect a difference in ongoing pregnancies leading to live births for three important subgroups of infertile women. Most randomized clinical trials regarding the fertility enhancing effect of oil-based contrast during HSG have been previously performed in couples with unexplained infertility and/or mild male factor. While several trials have also included women with other types of infertility (31-34), none were able to show a significant positive effect of tubal flushing with oil-based contrast mostly due to a low sample size, and the results of previous trials among women with unexplained infertility cannot simply be extrapolated to all women facing infertility (14). Another strength of this study is its multinational character. As it involves academic and non-academic (teaching and non-teaching) participating centres from the Netherlands and the United Kingdom, the results will be applicable to different countries with different hospital settings. A limitation of this study is the potential diversity in treatments between the various participating centres. The participating centres will treat patients according to their local protocol. Although these protocols are based on national guidelines, there is variety between the national guidelines of the participating countries (4, 35). As this study evaluates two variants of standard care which are already applied in current practice, we chose this pragmatic approach to generate evidence that is applicable to the majority of the treating centres. This approach will lead to a difference in management between participating centres, possibly influencing the chance of conception. Randomization is therefore stratified per inclusion group and per inclusion site, to prevent uneven distribution among the two randomization groups.

Potential impact and implications

As health care costs are increasing around the world, research focusing on cost-effectiveness of healthcare will help clinicians and policy makers to determine the appropriate position of the HSG with oil-based contrast within the fertility workup, taking into account both its diagnostic potential as well as its therapeutic effect. A cost-effectiveness analysis of the long term outcomes of the H2Oil study showed an increase in the cumulative pregnancy rate when oil-based contrast was used, compared to when water-based contrast was used (80,0 versus 75,0%) (36). The higher price of the oil-based contrast was compensated by a decrease in the need for ART to achieve these pregnancies in the group receiving oil-based contrast, resulting in comparable overall costs. The study concluded that tubal flushing with oil-based contrast was therefore cost-effective in comparison to water-based contrast in women with unexplained subfertility (36). In the proposed study, a cost-effectiveness analysis will be performed incorporating medical consumption, absence from (paid) work and loss of productivity due to health problems (24).

The mechanism of action of oil-based contrast leading to a fertility enhancing effect is not fully elucidated. Different hypotheses place the point of action in the Fallopian tube (37, 38), the endometrium, (39) and the peritoneum (40, 41). A post-hoc analysis of the H2Oil-study showed that in the group of women with higher pain scores, the ongoing pregnancy rate was higher in women that had received oil-based contrast during HSG when compared to women that had received water-based contrast (42). These result support the first hypothesis that when using oil-based contrast medium, the pain was caused by an increase in intrauterine pressure prior to dislodgment of pregnancy-hindering debris from the proximal part of otherwise anatomically normal Fallopian tubes (42). Previous research associated pre-procedural anxiety to a higher experienced pain level during medical procedures (19, 43). To further investigate the relationship between discomfort or pain during HSG and ongoing pregnancies in the current study, the APAIS questionnaire will be used to score pre-procedural anxiety as a confounder for experienced pain.

Despite reassuring recent research on the prevalence of complications after an HSG, fear of complications is still a reason for some clinicians to withhold use of oil-based contrast (11, 18). A recent review, without publication date or language restrictions, showed that the incidence of intravasation of contrast in the venous or lymphatic system is higher during tubal flushing with oil-based contrast in comparison to water-based contrast (2.8% and 1.8% respectively, odds ratio 5.05, 95% CI 2.27 to 11.22) in the included RCTs (18). However, when including only studies that used fluoroscopy screening during HSG, no serious consequences of intravasation were identified. Pelvic infection after HSG is

another well-known complication, the previously mentioned review described that in studies published in or after 1960, as antibiotics were not routinely used before, the frequency of infection was 0.55% after HSG with oil-based contrast and 0.35% after HSG with water-based contrast (18). The proposed study aims to provide information on the complication rate in a population of women that have a higher risk for tubal pathology (because of a previous infection, pelvic surgery or endometriosis). Complications will be reported systematically.

This multicentre RCT will establish whether, for women that are 39 years or older, women with ovulation disorders, and women who have a high risk for tubal pathology, there is a fertility enhancing effect of tubal flushing at HSG with oil-based contrast during fertility work-up and if this is cost-effective.

REFERENCES

1. Mascarenhas MN, Flaxman SR, Boerma T, Vanderpoel S, Stevens GA. National, regional, and global trends in infertility prevalence since 1990: a systematic analysis of 277 health surveys. PLoS Med. 2012;9(12):e1001356.
2. Zegers-Hochschild F, Adamson GD, Dyer S, Racowsky C, de Mouzon J, Sokol R, et al. The International Glossary on Infertility and Fertility Care, 2017. Hum Reprod. 2017;32(9):1786-801.
3. ACOG. Infertility Workup for the Women's Health Specialist. Obstetrics and Gynecology. 2019;133:377-84.
4. NICE. Fertility assessment and treatment for people with fertility problems. 2013 01-02-2013. Contract No.: CG 156 2013.
5. Mills M, Rindfuss RR, McDonald P, te Velde E, Reproduction E, Society Task F. Why do people postpone parenthood? Reasons and social policy incentives. Hum Reprod Update. 2011;17(6):848-60.
6. Hoenderboom BM, van Benthem BHB, van Bergen J, Dukers-Muijrers N, Gotz HM, Hoebe C, et al. Relation between Chlamydia trachomatis infection and pelvic inflammatory disease, ectopic pregnancy and tubal factor infertility in a Dutch cohort of women previously tested for chlamydia in a chlamydia screening trial. Sex Transm Infect. 2019;95(4):300-6.
7. Trimbos-Kemper T, Trimbos B, van Hall E. Etiological factors in tubal infertility. Fertility and Sterility. 1982;37(3):384-8.
8. Taylor E, Gomel V. The uterus and fertility. Fertil Steril. 2008;89(1):1-16.
9. Practice Committee of the American Society for Reproductive M. Effectiveness and treatment for unexplained infertility. Fertil Steril. 2006;86(5 Suppl 1):S111-4.
10. Bhattacharya S, Porter M, Amalraj E, Templeton A, Hamilton M, Lee AJ, et al. The epidemiology of infertility in the North East of Scotland. Hum Reprod. 2009;24(12):3096-107.
11. Roest I, van Welie N, Mijatovic V, Dreyer K, Bongers M, Koks C, et al. Complications after hysterosalpingography with oil- or water-based contrast: results of a nationwide survey. Hum Reprod Open. 2020;2020(1):hoz045.
12. Wang R, van Welie N, van Rijswijk J, Johnson NP, Norman RJ, Dreyer K, et al. Effectiveness on fertility outcome of tubal flushing with different contrast media: systematic review and network meta-analysis. Ultrasound Obstet Gynecol. 2019;54(2):172-81.
13. Fang F, Bai Y, Zhang Y, Faramand A. Oil-based versus water-based contrast for hysterosalpingography in infertile women: a systematic review and meta-analysis of randomized controlled trials. Fertil Steril. 2018(110):153-60.
14. Wang R, Watson A, Johnson N, Cheung K, Fitzgerald C, Mol BWJ, et al. Tubal flushing for subfertility. Cochrane Database Syst Rev. 2020;10:CD003718.
15. Dreyer K, van Rijswijk J, Mijatovic V, Goddijn M, Verhoeve HR, van Rooij IAJ, et al. Oil-Based or Water-Based Contrast for Hysterosalpingography in Infertile Women. N Engl J Med. 2017;376(21):2043-52.
16. van Rijswijk J, van Welie N, Dreyer K, Pham CT, Verhoeve HR, Hoek A, et al. Tubal flushing with oil-based or water-based contrast at hysterosalpingography for infertility: long-term reproductive outcomes of a randomized trial. Fertility and Sterility. 2020;114(1):155-62.
17. Lindequist S, Rasmussen F, Larsen C. Use of Iotrolan versus Ethiodized Poppy-Seed Oil in Hysterosalpingography. Radiology. 1994;191:513-7.
18. Roest I, Rosielle K, van Welie N, Dreyer K, Bongers M, Mijatovic V, et al. Safety of oil-based contrast medium for hysterosalpingography: a systematic review. Reproductive BioMedicine Online. 2021;42(6):1119-29.

19. Moerman N, van Dam FSAM, Muller MJ, Oosting H. The Amsterdam Preoperative Anxiety and Information Scale (APAIS). Anesth Analg. 1994;82:445-51.
20. Anovulation and wish to conceive: Dutch Society of Obstetrics and Gynaecology; 2004 [Available from: https://www.nvog.nl/wp-content/uploads/2017/12/Anovulatie-en-kinderwens-2.0-12-11-2004.pdf.
21. Basic fertility work-up: Dutch Society of Obstetrics and Gynaecology; 2015 [Available from: https://richtlijnendatabase.nl/richtlijn/orienterend_fertiliteitsonderzoek_ofo/startpagina_orienterend_fertilitietsonderzoek.html.
22. Unexplained infertility: Dutch Society for Obstetrics and Gynaecology; 2020 [Available from: https://richtlijnendatabase.nl/richtlijn/onverklaarde_subfertiliteit/startpagina_-_onverklaarde_subfertiliteit.html.
23. Hunault CC, Laven JS, van Rooij IA, Eijkemans MJ, te Velde ER, Habbema JD. Prospective validation of two models predicting pregnancy leading to live birth among untreated subfertile couples. Hum Reprod. 2005;20(6):1636-41.
24. Bouwmans C. Manual iMTA Productivity Cost Questionnaire (iPCQ) 2013 [Available from: www.imta.nl.
25. Kanters TA, Bouwmans CAM, van der Linden N, Tan SS, Hakkaart-van Roijen L. Update of the Dutch manual for costing studies in health care. PLoS One. 2017;12(11):e0187477.
26. NICE. The guidlines manual: Assessing cost effectiveness 2012 [Available from: https://www.nice.org.uk/article/pmg6/chapter/7-assessing-cost-effectiveness#economic-evidence-and-guideline-recommendations.
27. ICH. Good Clinical Practice (GCP) https://www.ich.org/: International Conference on Harmonisation of technical requirements for registration of pharmaceuticals for human use; 2016 [
28. WHO Trial Registration Data Set. Version 1.3.1: World Health Organization; [Available from: https://www.who.int/clinical-trials-registry-platform/network/who-data-set.
29. HFEA. Fertility treatment 2019: trends and figures 2021 [UK statistics for IVF and DI treatment, storage and donation]. Available from: https://www.hfea.gov.uk/about-us/publications/research-and-data/fertility-treatment-2019-trends-and-figures/.
30. SART. Preliminary National Summary Report 2019 2022 [Available from: https://www.sartcorsonline.com/rptCSR_PublicMultYear.aspx?ClinicPKID=0.
31. DeCherney AH, Kort H, Barney JB, DeVore GR. Increased pregnancy rate with oil-soluble hysterosalpingography dye. Fertil Steril. 1980;33(4):407-10.
32. Alper MM, Garner PR, Spence JEH, Quarrington AM. Pregnancy rates after hysterosalpingography with oil- and water-soluble contrast media. Obstetrics and Gynecology. 1986;68(1):6-9.
33. Schwabe MG, Shapiro SS, Haning RV. Hysterosalpingography with oil contrast medium enhances fertility in patients with infertility of unknown etiology. Fertil Steril. 1983;40(5):604-6.
34. Steiner AZ, Meyere WR, Clark RL, Hartmann KE. Oil-soluble contrast during hysterosalpingography in women with proven tubal patency. Obstetrics and Gynecology. 2003;101:109-13.
35. Infertility: Dutch Society of Obstetrics and Gynaecology; 2010 [Available from: https://www.nvog.nl/wp-content/uploads/2018/02/Subfertiliteit-landelijke-netwerkrichtlijn-1.0-20-05-2011.pdf.
36. van Welie N, Pham CT, van Rijswijk J, Dreyer K, Verhoeve HR, Hoek A, et al. The long-term costs and effects of tubal flushing with oil-based versus water-based contrast during hysterosalpingography. Reprod Biomed Online. 2021;42(1):150-7.

37. Kerin JF, Williams DB, San Roman GA, Pearlstone AC, Grundfest WS, Surrey ES. Falloposcopic classification and treatment of fallopian tube lumen disease. Fertility and Sterility. 1992;57(4):731-41.

38. Roest I. What is the fertility-enhancing effect of tubal flushing? A hypothesis article. International Journal of Obstetrics and Gynaecology2022.

39. Johnson NP, Bhattu S, Wagner A, Blake DA, Chamley LW. Lipiodol alters murine uterine dendritic cell populations: a potential mechanism for the fertility-enhancing effect of lipiodol. Fertil Steril. 2005;83(6):1814-21.

40. Johnson JV, Montoya IA, Olive DL. Ethiodol oil contrast medium inhibits macrophage phagocytosis and adherence by altering membrane electronegativity and microviscosity. Fertility and Sterility. 1992;58(3):511-7.

41. Mikulska D, Kurzawa R, Rozewicka L. Morphology of in vitro sperm phagocytosis by rat peritoneal macrophages under influence of oily contrast medium (Lipiodol). Acta Eur Fert. 1994.

42. van Welie N, Dreyer K, van Rijswijk J, Verhoeve HR, Goddijn M, Nap AW, et al. Treatment effect of oil-based contrast is related to experienced pain at HSG: a post-hoc analysis of the randomised H2Oil study. Hum Reprod. 2019;34(12):2391-8.

43. Sorrentino F, Petito A, Angioni S, D'Antonio F, Severo M, Solazzo MC, et al. Impact of anxiety levels on the perception of pain in patients undergoing office hysteroscopy. Arch Gynecol Obstet. 2021;303(4):999-1007.

SUPPLEMENTARY MATERIAL

Supplementary Table 1. Trial Registration Data

Acronym	H2Olie2
Title	H2Oil2: Oil- based versus water-based contrast media for hysterosalpingography (HSG) in infertile women with unevaluated indications: a randomized controlled trial
Scientific title	Oil- based versus water-based contrast media for hysterosalpingography (HSG) in infertile women with unevaluated indications: a randomized controlled trial
Summary	Rationale: We hypothesize that tubal flushing at hysterosalpingography (HSG) with oil-based contrast will result in higher pregnancy and live birth rates as compared to tubal flushing at HSG with water-based contrast in women: with an ovulation disorder, at high risk for tubal pathology and/or ≥38 years of age, which will lead to a reduction in the need for expensive fertility treatments like IVF and/or ICSI, and will therefore be a cost effective strategy. Objective: The objective of the proposed study is to assess the effectiveness and costeffectiveness of the use of oil versus water-based contrast medium in terms of live birth in women undergoing HSG, who: 1: have ovulation disorders and/or; 2: are at high risk for tubal pathology and/or; 3: are 39 years of age or over. Study design: Multicenter, randomized controlled trial with a cost-effectiveness analysis alongside it. Study population: We will study women: 1: with ovulation disorders and/or; 2: at high risk for tubal pathology and/or; 3: are 39 years of age or over. Intervention: We will compare tubal flushing at HSG with oil-based contrast (intervention) versus tubal flushing with water-based contrast (control). Main study parameters/endpoints: The primary outcome is conception leading to live birth, with a positive pregnancy test preceding the pregnancy within 6 months after randomization. We will also study time-to-pregnancy. Our hypothesis is that HSG with oil-based contrast will increase pregnancy. Nature and extent of the burden and risks associated with participation, benefit and group relatedness: As we compare strategies (HSG with oil-based contrast versus HSG with water-based contrast) that are already applied in current practice, no additional risks or burdens are expected from the study.
Status	Open for patient inclusion
Control group	Active
Grouping	Parallel
Arms	2 or more arms, randomized
Masking	None
Target size	930
Inclusion criteria	In order to be eligible to participate in this study, women must meet one of the following criteria: 1: with ovulation disorders (ovulation disorders will be defined as less than 8 menstrual cycles per year) or; 2: at high risk for tubal pathology (high risk for tubal pathology will be defined as a positive chlamydia infection, a pelvic inflammatory disease, known endometriosis, abdominal surgery (including tubectomy for ectopic pregnancy and appendectomy) and/or peritonitis in the medical history) or; 3: 39 years of age or over

Supplementary Table 1. (Continued)

Acronym	H2Olie2
Exclusion criteria	- Iodinated contrast agent allergy - Male subfertility defined as total motile sperm count < 3 x10^6 spermatozoa/ml - Not willing or able to sign the consent form
Start date	2019-08-20
Stop date	2023-08-01
Disease	Subfertility, tubal patency testing
Hypothesis	We hypothesize that tubal flushing at hysterosalpingography (HSG) with oil-based contrast will result in higher pregnancy and live birth rates as compared to tubal flushing at HSG with water-based contrast in the target population.
Interventions	HSG as tubal patency test with oil-based contrast versus HSG as tubal patency test with water-based contrast.
Primary outcome	Conception leading to live birth with a positive pregnancy test within 6 months after randomization.
Secondary outcome	- Biochemical pregnancy, clinical pregnancy, ongoin pregnancy - Miscarriage, ectopic pregnancy, multiple pregnancy - Time to pregnancy - Complications following HSG (infection, intravastion) - Pregnancy outcomes (f.e. birth weight) - Pregnancy complications - Stillbirth - Thyroid function of the woman (before and 1 month after HSG) - Neonatal outcomes - Additional fertility treatments (Intra-uterine insemination, IVF, IVF/ICSI) - Costs within 6 months after randomization - Thyroid function of neonate (determined by heelprick)
Sponsors	Amsterdam UMC, VUmc
Time points	Follow-up 6 months after randomization
MEC approved	Yes
Randomized	Yes
Plan to share IPD	Undecided
IPD plan description	N/A
Publications	N/A
Issueing body	METc VUmc
Source ID	NL66079.029.18
Funding sources	ZonMw, Guerbet
Old NTR ID	N/A
Date registered	2019-08-01
URL	www.H2Olie.nl

Supplementary file 2. List of currently participating centers H2Oil2 (March 1st 2022)

The Netherlands:

Amsterdam UMC location Vrije Universiteit, Amsterdam, PI Prof. dr. V. Mijatovic

Amstelland Ziekenhuis, Amstelveen, PI Drs. A. Mozes

Reinier de Graaf Groep, Delft, PI Dr. E.J.P. van Santbrink

Elkerliek Ziekenhuis, Helmond, PI dr. J. Penninx

Zaans Medisch Centrum, Zaandam, PI drs. A.B. Hooker

Catharina Ziekenhuis Eindhoven, PI dr. A.G. Huppelschoten

United Kingdom:

Imperial College NHS Healthcare Trust, London, PI prof. dr. A. Hemingway

CHAPTER 4

Does the effectiveness of IUI in couples with unexplained subfertility depend on their prognosis of natural conception? A replication of the H2Oil study

Rik van Eekelen, Kimmy Rosielle, Nienke van Welie, Kim Dreyer, Madelon van Wely, Ben Mol, Martinus Eijkemans, Velja Mijatovic, and Nan van Geloven

Human Reprod Open. 2020 Nov 17;2020(4)

ABSTRACT

Study question

Can we replicate the finding that the benefit of IUI-ovarian stimulation (IUI-OS) compared to expectant management for couples with unexplained subfertility depends on the prognosis of natural conception?

Summary answer

The estimated benefit of IUI-OS did not depend on the prognosis of natural conception but did depend on when treatment was started after diagnosis, with starting IUI-OS later yielding a larger absolute and relative benefit of treatment.

What is known already

IUI-OS is often the first-line treatment for couples with unexplained subfertility. Two randomized controlled trials (RCTs) compared IUI-OS to expectant management using different thresholds for the prognosis of natural conception as inclusion criteria and found different results. In a previous study (a Dutch national cohort), it was found that the benefit of IUI-OS compared to expectant management seemed dependent on the prognosis of natural conception, but this finding warrants replication.

Study design, size, duration

We conducted a secondary analysis of the H2Oil study (n=1119), a multicentre RCT that evaluated the effect of oil-based contrast versus water-based contrast during hysterosalpingography (HSG). Couples were randomized before HSG and followed up for 3–5 years. We selected couples with unexplained subfertility who received HSG and had follow-up or pregnancy data available. Follow-up was censored at the start of IVF, after the last IUI cycle or at last contact and was truncated at a maximum of 18 months after the fertility workup.

Participants/materials, setting, method

The endpoint was time to conception leading to an ongoing pregnancy. We used the sequential Cox approach comparing in each month the ongoing pregnancy rates over the next 6 months of couples who started IUI-OS to couples who did not. We calculated the prognosis of natural conception for individual couples, updated this over consecutive failed cycles and evaluated whether prognosis modified the effect of starting IUI-OS. We corrected for known predictors of conception using inverse probability weighting.

Main results and the role of chance

Data from 975 couples were available. There were 587 couples who received at least one IUI-OS cycle within 18 months after HSG of whom 221 conceived leading to an ongoing pregnancy (rate: 0.74 per couple per year over a median follow-up for IUI of 5 months). The median period between HSG and starting IUI-OS was 4 months. Out of 388 untreated couples, 299 conceived naturally (rate: 0.56 per couple per year over a median follow-up of 4 months). After creating our mimicked trial datasets, starting IUI-OS was associated with a higher chance of ongoing pregnancy by a pooled, overall hazard ratio of 1.50 (95% CI: 1.19–1.89) compared to expectant management. We did not find strong evidence that the effect of treatment was modified by a couple's prognosis of achieving natural conception (Akaike's Information Criterion (AIC) decreased by 1 point). The effect of treatment was dependent on when couples started IUI-OS (AIC decreased by more than 2 points). The patterns of estimated absolute chances over time for couples with increasingly better prognoses were different from the previous study but the finding that starting later yields a larger benefit of treatment was similar. We found IUI-OS increased the absolute chance of pregnancy by at least 5% compared to expectant management. The absolute chance of pregnancy after IUI-OS seems less variable between couples and starting times of treatment than the absolute chance after expectant management.

Limitations, reasons for caution

This is a secondary analysis, as the H2Oil trial was not designed with this research question in mind. Owing to sample size restrictions, it remained difficult to distinguish between the ranges of prognoses in which true benefit was found.

Wider implications of findings

We replicated the finding that starting IUI-OS later after diagnosis yields a larger absolute and relative benefit of treatment. We did not replicate the dependency of the effect of IUI-OS on the prognosis of natural conception and could not identify clear thresholds for the prognosis of natural conception when IUI-OS was and/or was not effective. Because many of these couples still have good chances of natural conception at the time of diagnosis, we suggest clinicians should advise couples to delay the start of IUI-OS for several months to avoid unnecessary treatment.

INTRODUCTION

Couples who have been trying to conceive for at least 12 months and whose fertility workup fails to reveal any abnormalities are considered to have unexplained subfertility (1, 2). In several countries intrauterine insemination (IUI) is used as first-line treatment in these couples, especially in combination with ovarian stimulation (OS), since IUI is less invasive and less costly than in vitro fertilisation (IVF) (3), despite the lack of evidence from randomized controlled trials (RCTs) regarding the effectiveness of IUI-OS (4). The two trials that compared IUI-OS to expectant management used different thresholds for the prognosis of natural conception as inclusion criteria (5, 6). In women with an intermediate prognosis to conceive naturally, i.e. an estimated probability between 30% and 40% to conceive within 12 months leading to live birth, IUI-OS was no more effective than expectant management (6). In women with a poor prognosis i.e. <30% over 12 months, IUI-OS did result in more live births than expectant management (5).

In a previous study, we found that the different outcomes of these two trials might be explained by the difference in the prognosis of natural conception (7). In a Dutch cohort of 1896 couples, we found that couples with lower prognoses of natural conception had more benefit from IUI-OS in terms of a relative and absolute difference in the chance of conception compared to expectant management. Due to sample size limitations, it was difficult to identify a fixed threshold for prognoses at which point IUI-OS becomes effective. For a prognosis below 25% over one year, IUI-OS seemed effective, leading to higher chances of ongoing pregnancy in six months compared to expectant management. For higher prognosis than 40% over one year, IUI-OS was not effective and led to similar chances of ongoing pregnancy compared to expectant management. Between these thresholds, it was uncertain whether IUI-OS was effective.

Replication of research findings is an essential part of medical research as many findings unfortunately cannot be reproduced in further studies (8). Addressing the same research question with different methodological approaches, such as trial and observational data, provides more evidence that a single result is not a chance finding (9, 10). More knowledge on who does and who does not benefit from IUI-OS can guide clinical practice and inform evidence-based shared decision making on when to start treatment. Because these thresholds hold great importance to patients, they should be based on solid evidence.

The aim of this study was to replicate the previous result (i.e. that the benefit of IUI-OS compared to expectant management for couples with unexplained subfertility depends on the prognosis of natural conception) in an independent data source derived from an

RCT on contrast fluid used for hysterosalpingography (HSG). This is a different approach to see if the effectiveness of IUI-OS here also depends on the prognosis of natural conception, as well as to validate the thresholds of 25% and 40%.

MATERIALS AND METHODS

The H2Oil study was a multicentre RCT to compare ongoing pregnancy rates in subfertile women who underwent HSG with oil-based versus water-based contrast (Netherlands Trial Register number, NTR3270). The study was approved by the Institutional Review Board of the Amsterdam University Medical Centre – Academic Medical Centre (reference 2008.362, dated 12 February 2009). The study details and results have been published previously (11).

The H2Oil follow-up study (NTR 6577) assessed the long term outcomes until 3–5 years after the H2Oil study and was approved by the Institutional Review Board of the Amsterdam University Medical Centre, location VU University Medical Centre (reference 2017.221, dated 14 June 2017). Study details and results have been published elsewhere (12).

Inclusion and exclusion criteria

Included women were between 18 and 39 years of age, were having ovulatory cycles and had a low perceived risk of tubal pathology based on their medical history. They had tried to conceive unsuccessfully for at least 1 year and had an indication for tubal patency testing. Exclusion criteria were known endocrine disorders and having a partner with severe male subfertility (defined as a total motile sperm count after sperm wash of less than 3 million per millilitre).

Women were randomized for an HSG with oil-based contrast or an HSG with water-based contrast. Data regarding fertility treatments and pregnancies were collected until 3–5 years after randomization.

Follow up and outcome definitions

For the follow up of selected couples, we distinguished between time spent pursuing expectant management and time spent receiving IUI-OS cycles. The start of the IUI period was defined as the first day of menstruation before the first IUI cycle. The end of the IUI period was defined as the first day of menstruation before the last IUI cycle. All pregnancies in the IUI period thus resulted from IUI. Follow up for expectant management started 14 days before they received HSG and ended at the last date of contact, first day

of last menstruation before starting IUI or IVF or, in case they conceived naturally, the first day of the last menstruation before conceiving.

The endpoint was ongoing pregnancy, defined as the presence of foetal cardiac activity at transvaginal sonography at a gestational age of at least 12 weeks (11). Couples who miscarried before 12 weeks were not censored since they could still achieve ongoing pregnancy in subsequent cycles after their miscarriage. If no ongoing pregnancy occurred, we censored follow up at the end of expectant management or, if treated, at the end of the IUI period.

Cumulative pregnancy rates over multiple IUI cycles

We used the same statistical approach as in our previous study (7). In short, we used the sequential Cox approach to compare multiple cycles of treated and untreated couples, not only directly after completion of the fertility workup but also if they started later (13).

In this approach, we derived multiple datasets from the cohort in which couples started IUI-OS at approximately the same point in time and compared them to couples undergoing expectant management at that time, 'mimicking' hypothetical RCTs (13). At completion of the fertility workup and each consecutive month thereafter, named the landmark time points, we constructed such a mimicked trial from our data in which we included all couples who remained in the cohort i.e. couples who had not conceived, had not started treatment and were not lost to follow up before that landmark time point. In these 'trial' datasets spanning 6 months, we considered couples as treated if they started IUI-OS early i.e. within one month after the landmark time point. Couples who did not start IUI-OS within the first month were used as controls. Couples who started IUI-OS within the 6 month window of a trial, but later than one month after the landmark time point, were counted as controls during their untreated period and 'artificially censored' at the time of starting IUI-OS. This way, couples were not included in a single group throughout the study. Instead, couples who at some point started IUI-OS were analysed as controls (under expectant management) in the 'mimicked' trial datasets preceding the month in which they started IUI-OS. When they started IUI-OS, their following treatment cycles were analysed as part of the treated (IUI-OS) group in the mimicked trial dataset that started that month.

In order to compare results to the previous study, we restricted our data to a maximum of eighteen months of follow up.

Adjusting for patient characteristics that differed between treated and untreated couple

In our data, couples were not randomized to either expectant management or IUI-OS. Thus, patients starting IUI could differ from those who did not in terms of important predictors of conception such as female age or duration of subfertility. In order to achieve groups that are on average similar, we opted for a statistical technique called applied iterative inverse probability weighting (14-16). By reweighting patients' contribution to the data, these characteristics are balanced. Details on how we derived the weights to adjust for these differences are given in the Supplementary Data. We chose to balance for the same patient characteristics as in the previous study with the exception of fertility clinic, as that would lead to very unbalanced weights: female age, duration of subfertility, primary or secondary subfertility, total motile sperm count, referral status and the presence of one-sided tubal pathology (17, 18). We calculated the mean weight to assess potential inflation of the effective sample size induced by the weighting, which is ideally around 1 (19). We assessed the degree of balance in patient characteristics before and after weighting using the standardized mean difference between the treated and untreated group in each of the mimicked trial dataset. A lower standardized mean difference between groups represents better balance and a value below 0.10 generally indicates no important difference (14, 15).

Statistical analysis

We analysed the weighted mimicked trial datasets using a pooled Cox proportional hazards model with IUI-OS or expectant management as a treatment covariate. We calculated an overall hazard ratio by stratifying on the 13 mimicked trials. We used a robust sandwich variance estimator to adjust precision measures since couples can be included in multiple mimicked trial datasets (20).

Modification of the estimated effect of IUI-OS by the prognosis of natural conception

To address whether the effect of starting IUI-OS depends on the decreasing prognosis of natural conception of the individual couple, we added the prognosis and a treatment-by-prognosis interaction term to the model. We calculated a time-updated prognosis of natural conception over the next 6 cycles at the start of each mimicked trial dataset by using an existing dynamic prediction model that comprises female age, duration of subfertility, primary or secondary subfertility, percentage of progressive motile sperm, referral by a general practitioner or specialist, and the unsuccessful number of menstrual cycles since the fertility workup (18). The prognosis for a couple that we used is thus not one fixed value throughout the study but decreases after consecutive failed natural cycles. We transformed the updated prognosis by taking the complementary log-log of

its value such that it is linear on the log-hazard scale used by the Cox model (21). We included the complementary log-log of this updated prognosis as a main effect, the main effect for treatment and the treatment-by-prognosis interaction effect in the pooled Cox model. The weighting procedure was adjusted slightly for this analysis because the difference in prognosis between groups was adjusted for by adding it to the model as a main effect (22), (see also the Supplementary Materials).

For three hypothetical couples, we visually depicted the relationship between their worsening prognoses and the accompanying 6-month cumulative predicted probability of conception following expectant management or starting IUI-OS, as treatment is initiated later. The first example is a couple referred by their general practitioner, where the female partner is nulliparous and 32 years old, the couple has 1 year of subfertility at the time of completion of the fertility workup and the semen analysis showed 37% progressively motile sperm. In this case, the estimated prognosis of natural conception over the first 6 cycles is 25%. A second couple with the same characteristics except for a 2 year duration of subfertility at the completion of the fertility workup has a prognosis of 20% while a third couple with the same characteristics but for a 3.5-year duration of subfertility has a prognosis of 15%. At the time of the completion of their fertility workup, these couples have prognoses of 25%, 20% and 15% respectively over 6 cycles, which translates to approximately 40%, 32% and 25% respectively over 13 cycles i.e. one year (18).

The chances of natural conception for these three hypothetical couples decrease over time based on the number of unsuccessful menstrual cycles between the diagnosis/HSG and the start of a landmark.

Estimated cumulative probabilities of ongoing pregnancy from this model are derived from the separate mimicked trials that all have different observed conception rates, thus although predictions are expected to decrease over time, our estimates may fluctuate. We considered an absolute difference of more than five percentage points between estimates of the cumulative ongoing pregnancy rates, estimated at completion of the fertility workup, to indicate a benefit of IUI-OS.

In addition to modelling the impact of prognosis and consecutive failed natural cycles on the effect of treatment, we assessed if the effect of IUI-OS depends on the time of initiation of treatment by adding an additional interaction between treatment and landmark time point to the pooled Cox model already including treatment, prognosis and the treatment-by-prognosis interaction. If the interaction between prognosis and treatment yielded a better fit, we also added a three-way interaction between treatment,

prognosis and landmark time point to the previous model to see if the effect modification of prognosis on IUI-OS changed over mimicked trials i.e. when starting treatment later.

We used Akaike's Information Criterion (AIC) (at least two points difference) and Wald tests for the interaction terms to determine whether including the interactions resulted in a better fit of the model to the data (23).

Missing data

Data were missing on duration of subfertility (n=3), referral status (n=2), primary or secondary subfertility (n=1), the percentage of progressive motile sperm (n=81) and total motile sperm count (n=93) and was accounted for using single imputation. All statistical analyses were performed using R version 3.3.2 24)(*R Core Team (2017). http://www.R-project.org/.*) using the *survival*, *dynpred*, *mice* and *CreateTableOne* packages.

RESULTS

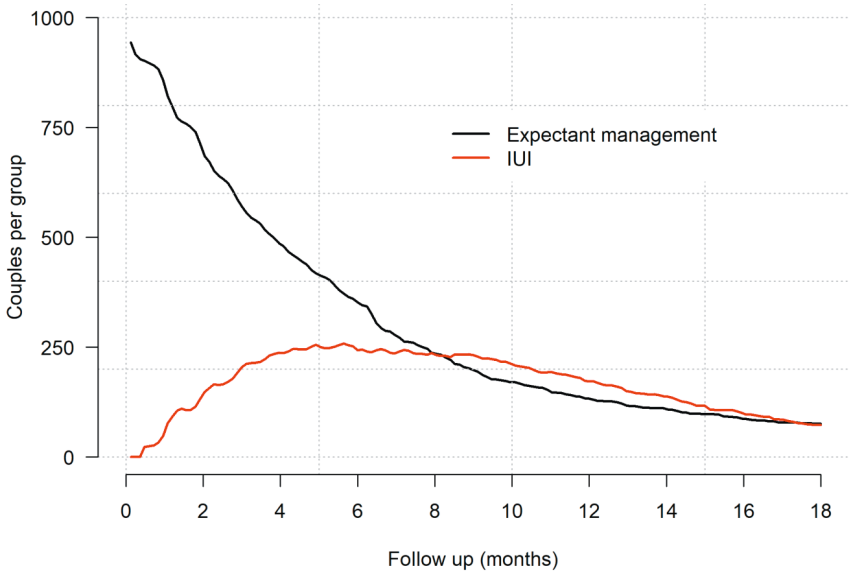
Out of 1119 couples included in the H2Oil trial, we selected 975 for analysis after excluding couples with other diagnoses than unexplained subfertility, couples who conceived before HSG and couples with missing outcome data.

Of these 975, 587 couples (60%) received 2386 IUI cycles after HSG, of whom 221 couples conceived after IUI, leading to ongoing pregnancy (rate: 0.74 per couple per year over a median follow up for IUI of 5 months). Out of 388 couples (40%) followed up for one and a half years of expectant management after HSG, 299 conceived naturally leading to ongoing pregnancy (rate: 0.56 per couple per year over a median follow-up of 4 months).

In total, 62 couples out of 587 (11%) who underwent IUI-OS started treatment directly after HSG and the remaining 525 (89%) first had a period of expectant management. The median period between HSG and starting IUI-OS was 4 months. A total of 1723 (72%) IUI cycles used OS. Forty-two couples (4%) received IVF as their first treatment, with a median period of expectant management of nine months between completion of the fertility workup and the start of IVF.

We depicted the number of couples followed under expectant management or followed under IUI-OS over time in Figure 1. Until approximately half a year of follow up the number of couples who were currently in an IUI-OS treatment pathway kept increasing, after which this number declined again.

Figure 1. Number of couples who, over follow up, are currently on expectant management or receiving IUI-OS



The baseline characteristics for couples who eventually received at least one cycle of IUI-OS within one and a half years after HSG or who remained untreated are summarized in Table 1. Treated couples more often had primary subfertility (73% vs. 60%) compared to couples that were not treated. Female age, median duration of subfertility, total motile sperm count, one-sided tubal pathology and referral status were similar between groups.

In the 13 weighted mimicked trial datasets, the standardized mean differences between treated and untreated couples were below 0.10 for all characteristics, indicating that the two groups were well balanced in terms of prognostic factors after weighting. The mean weight used in the pooled dataset was 1.00, indicating that weights are stable and do not artificially inflate sample size.

Effect estimates of IUI-OS

Starting IUI-OS was associated with increased ongoing pregnancy rates compared to expectant management as shown by an estimated hazard ratio of 1.50 (95%CI: 1.19–1.89), pooling all 13 mimicked, weighted trial datasets running over 6 months.

Tabel 1. Baseline characteristics of patients just before receiving HSG (n=975). Data are means (unless specified) or n (%) with the5th-95th percentile in brackets

	Couples who remained on expectant management (n=388)	**Couples who started IUI-OS within one and a half years (n=587)**
Female age (years)	32.7 (26.2-38.9)	32.9 (26.1-39.0)
Duration of subfertility (years, median)	1.6 (0.9-4.3)	1.7 (0.9-4.0)
Primary subfertility (versus secondary)	233 (60%)	426 (73%)
Total motile sperm cell count (millions, median)	72 (4-304)	43 (4-294)
One sided tubal pathology[a] (yes versus no)	16 (4%)	29 (5%)
Referral by specialist (versus referral by general practitioner)	31 (8%)	60 (10%)

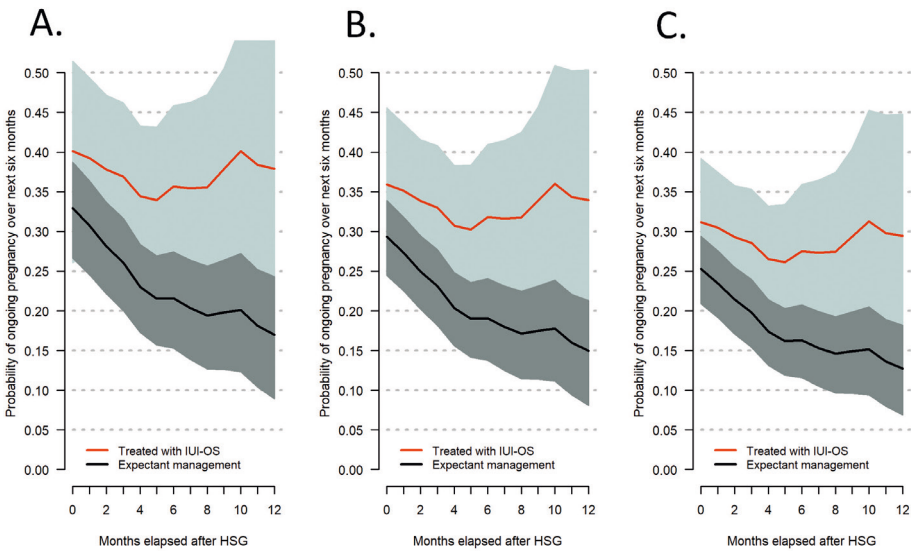
[a] assessed by hysterosalpingography (HSG) or in addition, a later laparoscopy

The predicted probability that a couple would conceive over the course of 6 months of expectant management after HSG was 29% (95% CI: 25–32%). If the couple started IUI-OS directly after HSG, their estimated probability of conception in the next six months was 40% (32–47%).

Judging by a decrease of at least 2 points in terms of AIC, the relative effect of IUI-OS did not depend on the prognosis of natural conception (AIC decreased by 1 point, p=0.17). The relative effect of IUI-OS was dependent on how long after the HSG treatment it was started (AIC decreased by 2.8 points, p=0.08).

The relations between prognosis, the start of treatment and the estimated treatment effect are visualized in Figure 2. The figure shows the 6-month cumulative probabilities of conception with and without starting IUI-OS for three different example couples with a prognosis to conceive naturally at completion of the fertility workup over the next year of 40% (Figure 2, left panel), 32% (Figure 2, middle panel) or 25% (Figure 2, right panel), which were updated over time when these couples fail consecutive natural cycles and start treatment later.

Figure 2. The association between the predicted prognosis of natural conception and the estimated benefit of starting IUI-OS at different time points.



The association between the predicted prognosis of natural conception and the estimated benefit of starting IUI-OS at different time points. This association is shown as cumulative probabilities over 6 months (y-axis) when starting IUI-OS, or not, at different time points after completion of the fertility workup (x-axis) for three example couples that have three different prognoses at time of hysterosalpingography (HSG): 40% (left), 32% (middle) or 25% (right). The prognosis was calculated over 1 year and updated after additional failed natural cycles. Grey bands represent 95% CIs. Left panel: Couple A is referred by their general practitioner, where the female partner is nulliparous and 32 years old, the couple has 1 year of subfertility at the time of completion of the fertility workup and the semen analysis showed 37% progressively motile sperm. Middle panel: Couple B has the same characteristics as Couple A except for 2 year duration of subfertility at the completion of the fertility workup. Right panel: Couple C has the same characteristics as Couple A but for 3.5 year duration of subfertility.

In Figure 2, the absolute chance to conceive over 6 months decreased over time for expectant management, but not for IUI-OS, of which the absolute chance seemed much less variable between couples and timing of treatment start, at around 37%, 34% and 30% for the three couples. However, CIs were wide, especially for pregnancy chances after IUI-OS. The decrease in chances for expectant management over time led to a larger treatment benefit as IUI-OS was started later.

It follows from Figure 2 that the prognosis for a couple does not have a large influence on the expected benefit in terms of the absolute difference between the IUI-OS and expectant management line as there was always a difference of 5% or more. This

was different in the previous study, in which the benefit of IUI-OS was dependent on prognosis. However, the later that treatment was started, the larger the expected benefit of treatment. This was the same finding as in the previous study.

DISCUSSION

We replicated the finding that in couples with unexplained subfertility, starting IUI-OS within one and a half years after completion of the fertility workup was associated with increased ongoing pregnancy rates over 6 months compared to expectant management. However, the estimated benefit of treatment did not depend on the prognosis of natural conception but did depend on when treatment was started after diagnosis. We replicated the finding that starting IUI-OS later yields a larger absolute and relative benefit of treatment.

The main strength of this study was the use of trial data from the H2Oil study with a follow-up of 3–5 years, low loss-to-follow up and few missing data. For the purpose of triangulation i.e. the use of multiple approaches to address the same question (9, 10), we now have data from two RCTs comparing IUI-OS to expectant management in different patient selection, cohort data and data from a RCT in which couples could receive IUI-OS during follow up. Workup and treatment protocols differed between these three data sources. Via triangulation, our confidence in two findings has been strengthened: namely that starting IUI-OS later yields a larger absolute and relative benefit of treatment and that the absolute chance of an ongoing pregnancy after IUI-OS is less variable between couples than the chance of natural conception. We did not find significant evidence that the effect of IUI-OS depends on the prognosis for natural conception of a couple.

Weaknesses are that the H2Oil trial was not designed with this secondary question in mind. In addition, the sample size was moderate which led to wide CIs and perhaps limited the power to show an interaction with prognosis.

We observed that the chances of an ongoing pregnancy after 6 months of expectant management in the present study were much higher than what was found in previous studies, at 29% instead of approximately 18% (17, 27). This could be due to the fact that in the H2Oil study, all couples received HSG during the diagnostic work-up, which might increase their chances, especially when using oil-based contrast medium (11, 26).

The pooled, i.e. overall effect of IUI-OS versus expectant management was less strong in the current study, with a point estimate for the hazard ratio of 1.50 compared to 1.96 that we found previously. The interaction that showed the dependency of the effect of IUI-OS on the prognosis for natural conception was in the same direction as the

previous study, with lower prognoses having more benefit of IUIOS, but this did not reach statistical significance in terms of AIC or P-value. It could be that we lacked the power, owing to a less strong main treatment effect and sample size restrictions, to show this dependency as we did in the previous study, which had a larger sample size of 1896 with 800 couples treated with IUI (7).

It remains unknown why couples with better prognoses would benefit less from IUI-OS. A possible mechanism is that there is a 'ceiling' to the chance of conception for subfertile couples in terms of a maximum and that some unexplained subfertile couples with good prognoses remain around this ceiling whereas couples with different indications are further below their fertility potential, which is increased by IUI-OS (28).

Not all couples who received IUI-OS conceived or continued IUI-OS over the 6 months follow up of a 'mimicked' trial dataset. In the previous study, there was no follow-up of natural conception after IUI-OS dropout. We repeated this approach in the current study.

In contrast to the previous study, we found that the effect of IUI-OS depended on when treatment was started independent of the decreasing prognosis of natural conception after failed cycles. This might be due to additional selection over time that is not explained by the dynamic prediction model (7, 18, 25). We found that when accounting for this time effect, the absolute chances after IUI-OS were much more stable over time than chances after expectant management, as the latter clearly decreased over time whereas the former did not. This provides more evidence that the chance after IUI-OS is less dependent on individual factors i.e. that couples' chances become more similar when receiving IUI-OS. This can be important for counselling couples, as it suggests that further expectant management will not come at a great loss in terms of a decreased chance of pregnancy when receiving IUI-OS later.

CONCLUSION

We replicated the finding that on average, IUI-OS increases the chance of an ongoing pregnancy compared to expectant management and that when IUI-OS is started later, the expected benefit in terms of the absolute and relative difference with expectant management was larger. We did not replicate the finding that the benefit of IUI-OS depends on the prognosis of natural conception. Couples with unexplained subfertility still have good chances of natural conception at the time of diagnosis and treatment is thus not always necessary. Clinicians should counsel couples on the option to prolong expectant management before commencing with IUI-OS.

REFERENCES

1. Aboulghar M, Baird D, Collin J, Evers J, Fauser B, Lambalk C, et al. Intrauterine insemination. Human reproduction update. 2009;15(3):265-77.
2. Brandes M, Hamilton CJ, de Bruin JP, Nelen WL, Kremer JA. The relative contribution of IVF to the total ongoing pregnancy rate in a subfertile cohort. Human reproduction (Oxford, England). 2010;25(1):118-26.
3. Tjon-Kon-Fat RI, Bensdorp AJ, Bossuyt PM, Koks C, Oosterhuis GJ, Hoek A, et al. Is IVF-served two different ways-more cost-effective than IUI with controlled ovarian hyperstimulation? Hum Reprod. 2015;30(10):2331-9.
4. Wang R, Danhof NA, Tjon-Kon-Fat RI, Eijkemans MJ, Bossuyt PM, Mochtar MH, et al. Interventions for unexplained infertility: a systematic review and network meta-analysis. The Cochrane database of systematic reviews. 2019;9:Cd012692.
5. Farquhar CM, Liu E, Armstrong S, Arroll N, Lensen S, Brown J. Intrauterine insemination with ovarian stimulation versus expectant management for unexplained infertility (TUI): a pragmatic, open-label, randomised, controlled, two-centre trial. Lancet. 2018;391(10119):441-50.
6. Steures P, van der Steeg JW, Hompes PG, Habbema JD, Eijkemans MJ, Broekmans FJ, et al. Intrauterine insemination with controlled ovarian hyperstimulation versus expectant management for couples with unexplained subfertility and an intermediate prognosis: a randomised clinical trial. Lancet. 2006;368(9531):216-21.
7. van Eekelen R, van Geloven N, van Wely M, McLernon DJ, Mol F, Custers IM, et al. Is IUI with ovarian stimulation effective in couples with unexplained subfertility? Human reproduction (Oxford, England). 2019a;34(1):84-91.
8. Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. Jama. 2005;294(2):218-28.
9. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. International journal of epidemiology. 2016;45(6):1866-86.
10. Munafò MR, Smith GD. Repeating experiments is not enough. Nature. 2018;553(7689):399-401.
11. Dreyer K, van Rijswijk J, Mijatovic V, Goddijn M, Verhoeve HR, van Rooij IAJ, et al. Oil-Based or Water-Based Contrast for Hysterosalpingography in Infertile Women. N Engl J Med. 2017;376(21):2043-52.
12. van Rijswijk J, van Welie N, Dreyer K, Pham CT, Verhoeve HR, Hoek A, et al. Tubal flushing with oil- or water-based contrast at hysterosalpingography for infertility: long-term reproductive outcomes of a randomized trial. Fertil Steril (accepted). 2020.
13. Gran JM, Roysland K, Wolbers M, Didelez V, Sterne JA, Ledergerber B, et al. A sequential Cox approach for estimating the causal effect of treatment in the presence of time-dependent confounding applied to data from the Swiss HIV Cohort Study. Statistics in medicine. 2010;29(26):2757-68.
14. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate behavioral research. 2011;46(3):399-424.
15. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Statistics in medicine. 2015;34(28):3661-79.
16. van der Wal W. Causal modelling in epidemiological practice [PhD thesis, 2011]. Chapter 8: Using iterative probability weighting to improve causal effect estimates. University of Amsterdam, Amsterdam, the Netherlands. Available from: http://bit.ly/2kWMrRt 2011.

17. Hunault CC, Habbema JD, Eijkemans MJ, Collins JA, Evers JL, te Velde ER. Two new prediction rules for spontaneous pregnancy leading to live birth among subfertile couples, based on the synthesis of three previous models. Hum Reprod. 2004;19(9):2019-26.
18. van Eekelen R, Scholten I, Tjon-Kon-Fat RI, van der Steeg JW, Steures P, Hompes P, et al. Natural conception: repeated predictions over time. Human reproduction (Oxford, England). 2017a;32(2):346-53.
19. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. American journal of epidemiology. 2008;168(6):656-64.
20. Wei L, Lin D, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. Journal of the American Statistical Association 1989;84:1065-73.
21. Prentice RL, Gloeckler LA. Regression analysis of grouped survival data with application to breast cancer data. Biometrics. 1978;34(1):57-67.
22. VanderWeele TJ. On the distinction between interaction and effect modification. Epidemiology (Cambridge, Mass). 2009;20(6):863-71.
23. Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 1974;19(6):716-23.
24. R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.
25. van Eekelen R, van Geloven N, van Wely M, McLernon DJ, Eijkemans MJ, Repping S, et al. Constructing the crystal ball: how to get reliable prognostic information for the management of subfertile couples. Human reproduction (Oxford, England). 2017b;32(11):2153-8.
26. Dreyer K, van Eekelen R, Tjon-Kon-Fat RI, van der Steeg JW, Steures P, Eijkemans M, et al. The therapeutic effect of hysterosalpingography in couples with unexplained subfertility: a post-hoc analysis of a prospective multi-centre cohort study. Reprod Biomed Online 2019;38:233–239.
27. van der Steeg JW, Steures P, Eijkemans MJ, Habbema JD, Hompes PG, Broekmans FJ, et al. Pregnancy is predictable: a large-scale prospective external validation of the prediction of spontaneous pregnancy in subfertile couples. Hum Reprod 2007;22:536–542.
28. Moreau J, Gatimel N, Simon C, Cohade C, Lesourd F, Parinaud J, Le´andri R. Potential chances for natural fertility influence results of intrauterine inseminations. Eur J Obstet Gynecol Reprod Biol X 2019; 4:100058.

SUPPLEMENTARY MATERIALS

The supplementary data to this manuscript can be found online at Human Reproduction Open.

CHAPTER 5

Safety of oil-based contrast medium for hysterosalpingography: a systematic review

Inez Roest, Kimmy Rosielle, Nienke van Welie, Kim Dreyer, Marlies Bongers, Velja Mijatovic, Ben W Mol, Carolien Koks

Reprod Biomed Online. 2021 Jun;42(6):1119-1129

ABSTRACT

Recent meta-analyses have shown that a hysterosalpingography (HSG) with oil-based contrast increases pregnancy rates in subfertile women. However, the frequency of complications during or after an HSG with oil-based contrast in subfertile women and/or their offspring is still unclear. This systematic review and meta-analysis, without restrictions on language, publication date or study design, was performed to fill this knowledge gap. The results show that the most frequently reported complication was intravasation of contrast, which occurred in 2.7% with the use of oil-based contrast (31 cohort studies and randomized controlled trials [RCT], 95% CI 1.7–3.8, absolute event rate 664/19,339), compared with 2.0% with the use of water-based contrast (8 cohort studies and RCT, 95% CI 1.2–3.0, absolute event rate 18/1006). In the cohort studies and RCT there were 18 women with an oil embolism (18/19,339 HSG), all without serious lasting consequences. Four cases with serious consequences of an oil embolism were described (retinal oil embolism [n = 1] and cerebral complaints [n = 3]); these reports did not describe the use of adequate fluoroscopy guidance during HSG. In conclusion, the most frequently reported complication after an HSG with oil-based contrast is intravasation occurring in 2.7%. In total four cases with serious consequences of oil embolisms in subfertile women were published.

Key message

The most frequently reported complication after an HSG with oil-based contrast is intravasation, occurring in 2.7% of HSG procedures. In total only four cases with serious consequences of oil embolisms in subfertile women were published. Therefore, safety concerns should not be the reason to deny the use of oil-based contrast for tubal testing in women with unexplained subfertility.

INTRODUCTION

Hysterosalpingography (HSG) to assess tubal patency is an essential part of the workup for subfertile couples (1). The first HSG was performed in 1910 by Rindfleisch (2). From 1914 iodized oils were used as an alternative to the water-based contrasts, which were irritative to the peritoneum (3, 4, 5).Different iodized oils were introduced, such as Lipiodol®, Iodochlorol, Ethiodol, Jodipin, Jodumbrin and Lipiodol® Ultra Fluid. The oil-based contrasts available today are Lipiodol® Ultra Fluid (Guerbet, Villepinte, France) and Ethiodized Poppyseed Oil (Heng Rui Pharmaceuticals, Jiangsu, China), the latter being currently only available in Asia.

Lipiodol was developed in 1901 as a solution containing iodine, and was used for a wide range of indications, including the reduction of struma and infection prevention. After the discovery of its radiological qualities, it was used for visualization of the uterine cavity and Fallopian tubes, but also in myelography, bronchography and later in lymphography. In 1960 a transesterified version of Lipiodol was developed, Lipiodol Ultra Fluid, which had a lower viscosity (6, 7).

For nearly seven decades, the therapeutic effect of oil-based contrast during HSG in the fertility work-up has been debated. Recently two meta-analyses have shown a favourable effect of oil-based contrast on fertility outcomes, with an OR of 1.47 (95% CI 1.12–1.93) for ongoing pregnancy and 2.18 (95% CI 1.30–3.65) for live birth when comparing HSG with oil-based contrast to water-based contrast (8, 9). This generated a worldwide renewed interest in the use of oil-based contrast for fertility enhancement. However, some clinicians are still hesitant about its use because of complications that have been reported in the past.

In 1929 the first report of intravasation of oil-based contrast during HSG was published (10). Intravasation is the inflow of contrast in the venous or lymphatic system, and is visualized by radiography, ideally with the use of fluoroscopy screening. Even though water-based contrast can also intravasate, only oil-based contrast is known to enter the circulation as droplets because of its hydrophobic qualities. These oil droplets can reach organs such as the lungs or brain as oil emboli and cause inflammation and/or occlusion of the vasculature (11). After this first case, more reports of intravasation followed, but most patients had only minor symptoms and recovered after observation. Intravasation was therefore regarded as innocuous (5). Currently, intravasation with the use of oil-based contrast is estimated to occur in around 5% of the HSG in the Netherlands (12).

In spite of this, a recent case report describes a patient falling into a comatose state as a result of an oil embolus after HSG (11). Although this might be a rare complication, it does emphasize the importance of safety and knowledge of the complication rates after HSG with the use of oil-based contrast.

As previously mentioned, Lipiodol contains iodine, and the iodine concentration in Lipiodol is higher than in water-based contrast (480 mg iodine/ ml in Lipiodol versus 240–300 mg iodine/ml in water-based contrast). Iodine exposure can cause a transient decrease in the synthesis of thyroid hormone (13). Subclinical hypothyroidism is associated with pregnancy complications (14). Furthermore, the HSG procedure has a risk of infection.

The systematic reviews and meta-analyses to date have primarily focused on fertility outcomes and have excluded case reports. This systematic review and meta-analysis included all study types, to provide an overview of the frequency and clinical consequences of all possible complications during or after HSG with the use of oil-based contrast in subfertile women.

MATERIALS AND METHODS

The protocol of this review was prospectively registered on PROSPERO (https://www.crd.york.ac.uk/prospero/, registration ID: CRD42018102382, registration date: 24 July 2018). The methodology used was as described in the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement (15).

Information sources and search strategies

Electronic databases including MEDLINE, EMBASE and the Cochrane Central Register of Controlled Trials (CENTRAL) were searched up to June 2020. Textbooks as well as reference lists of identified publications were also manually screened. The key search items included ‘hysterosalpingography’, ‘oil contrast’, ‘ethiodized oil’, ‘ethiodol’, ‘lipiodol’, ‘adverse effect’, ‘side effect’, ‘complication’, ‘thyroid’, ‘intravasation’, ‘embolization’, ‘granuloma’, ‘anaphylaxis’, ‘pelvic inflammatory disease’, and ‘adnexitis’ (Supplementary Tables 1–3).

Eligibility criteria

All types of studies were included: randomized controlled trials (RCT), prospective and retrospective cohort studies, case series and case reports that report complications occurring during or after HSG with the use of oil-based contrast, with or without comparison to water-based contrast, in women trying to conceive or their offspring. No

limitations on language or publication period were applied. Colleagues who were fluent in the foreign languages assisted in translating.

Outcomes

The outcomes included adverse events of HSG with the use of oil-based contrast (versus water-based contrast) in subfertile women and their offspring, such as: intravasation of the contrast medium, embolization of the contrast medium, pelvic inflammatory disease, lipogranuloma formation, retention of contrast, maternal or fetal thyroid dysfunction, and anaphylactic reactions. The clinical consequences included additional treatments, hospital stay, morbidity and mortality.

Study selection, data collection and quality assessment

Study eligibility was evaluated by two reviewers (IR and KR) independently; disagreements between the two reviewers were solved by consensus or by consultation with another reviewer (CK) when necessary. A predesigned form was used to extract the data and assess the quality of the included studies.

The following information was collected: name of the first author, publication year, study design, study population, participants' characteristics, types of contrast, details of interventions and co-interventions, sample sizes and outcomes. Full-text articles of English cohort and randomized studies were screened by a second reviewer (KR).

Risk of bias was assessed for all studies, excluding the case reports/series, in accordance with the quality assessment checklist for prevalence studies (16) (Supplementary Table 4). This checklist contains nine questions, each scored with 0 or 1 points. A total of 0–3 points is classified as an overall low risk of study bias, 4–6 points as moderate risk and 7–9 points as high risk. The risk of bias was assessed by two reviewers independently for the English studies.

Statistical analysis

The prevalence of complications occurring with the use of oil-based contrast was calculated, and where possible comparisons were made to the use of water-based contrast. Meta-analyses were performed using Review Manager Version 5.3. Statistical heterogeneity was estimated by performing a chi-squared test and calculating I2. Pooled weighted prevalences and the 95% CI were calculated using the MetaXL tool (Version 5.3, 2016; EpiGear International Pty Ltd, Queensland, Australia). A non-pre-specified sensitivity analysis was performed, selecting the cohorts and RCT to calculate the prevalence of complications. Case reports and case series were included to report all (and rare) complications.

RESULTS

Characteristics of included studies

The search identified 492 records. A total of 8 RCT, 41 cohort studies (4 prospective cohorts, 24 retrospective cohorts, 13 cohort studies which were not further specified) and 59 case reports/case series were included within the review. In these studies, a total of 23,536 HSG procedures were performed with the use of oil-based contrast (23,298 HSG in cohort studies/RCT). Sixteen of the included studies reported on HSG with water-based contrast as well, with a total of 1,975 HSG with water-based contrast (1,973 HSG in cohort studies/RCT) (for flow chart see Supplementary Figure 1). The included studies were published between 1928 and 2020 (see Supplementary Table 5 for the characteristics of the included studies) (10, 11, 17-121).

Quality of evidence of the studies

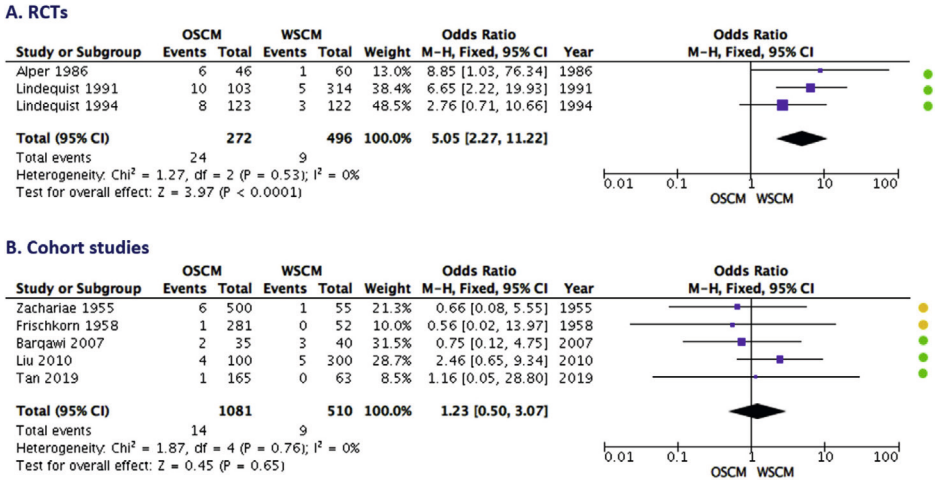
Of the 49 cohort studies and RCT, 16 studies were classified as low risk, 31 studies as moderate risk and two studies as high risk of study bias. In 18 studies, there was no clear definition of the reported complications. Mainly, there was no predefined definition of intravasation or oil embolism. There is no reliable or valid classification method for intravasation, therefore 44 of the 48 studies were classified as high risk of bias for the reliability and validity of the study instrument that measured the parameter of interest (see Supplementary Table 6 for the classification of all studies).

Intravasation and oil embolisms

Eight studies (three RCT and five cohort studies) compared the frequency of intravasation between HSG with the use of oil-based and water-based contrast (Figure 1). Rates of intravasation were 2.8% (38/1353) after HSG with oil-based contrast and 1.8% (18/1006) after HSG with water-based contrast (OR 5.05; 95% CI 2.27–11.22; $P < 0.0001$) based on the RCT and 1.23 (95% CI 0.50–3.07; $P = 0.65$) based on the cohort studies), showing that intravasation occurs more frequently with the use of oil-based contrast.

Twenty-three additional cohort studies reported on the prevalence of intravasation with the use of oil-based contrast alone. The overall pooled weighted frequency of intravasation in the 31 RCT and cohort studies with the use of oil-based contrast was 2.7% (95% CI 1.7–3.8, absolute event rate 664/19,339), compared with 2.0% (95% CI 1.2–3.0, absolute event rate 18/1,006) in the eight studies with the use of water based contrast. When including only studies published from 2000 onwards, the pooled frequency of intravasation with the use of oil-based contrast was 2.8% (95% CI 1.2–5.1, absolute event rate 12/471), compared with 1.8% (95% CI 0.0–5.9, absolute event rate 8/403) with the use of water-based contrast.

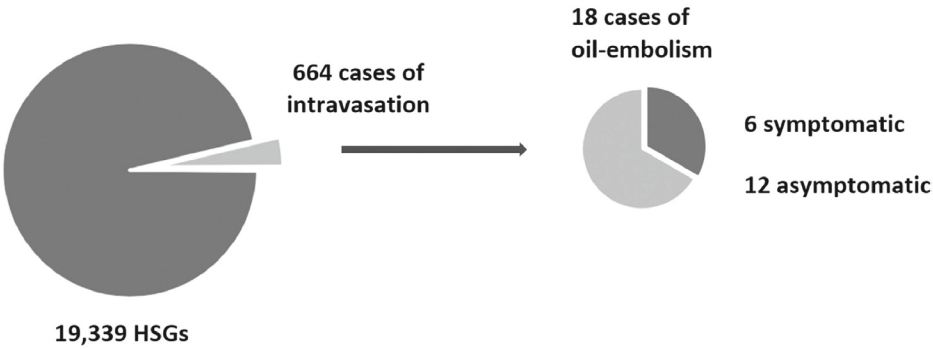
Figure 1. Prevalence of intravasation of oil-based contrast versus water-based contrast in HSG for subfertility.



Forest plot of meta-analysis reporting on intravasation with the use of oil-based contrast compared with water-based contrast. (A) RCT. (B) Cohort studies. OR and 95% CI. OR < 1 favour oil-based contrast (fewer adverse events); OR > 1 favour water-based contrast (fewer adverse events). The risk of bias of individual studies is represented by coloured dots: green (low risk of bias) and yellow (moderate risk of bias). HSG = hysterosalpingography; OSCM = oil-based contrast media; RCT = randomized controlled trial; WSCM = water-based contrast media.

In the whole group of HSGs with the use of oil-based contrast performed in RCT and cohort studies, there were 18 women with oil embolisms (18/19,339, 0.1% of HSG; 18/664, 2.7% of cases with intravasation). In six of these cases pulmonary embolisms were described, while the other 12 cases only described the contrast moving rapidly out of the pelvis. The latter were all asymptomatic and serious lasting consequences were not reported (see Figure 2).

Figure 2. Intravasation and oil embolisms in HSG with oil-based contrast for subfertility in cohort studies and RCT. HSG = hysterosalpingography; RCT = randomized controlled trial.



Additionally, there were 197 cases of intravasation after an HSG with the use of oil-based contrast in the case reports/series. In 22 of these women this led to the formation of an oil embolism (22/197, 11.2%). Four of these women were asymptomatic, 18 were symptomatic. Symptoms included a transient cough and/or dyspnoea and neurological symptoms. Four cases were described of women with serious consequences of an oil embolism (Table 1) (11, 30, 33, 41).

Table 1. Characteristics of serious consequences of oil embolism after HSG

Study	Contrast	Risk factors	Organ system involved	Consequences
Flew, 1944 (30)	Lipiodol (not specified)	HSG on day 24 of menstrual cycle; use of fluoroscopy not reported	Pulmonary and cerebrum	Hemiplegia, survived
Charawanamuttu et al, 1973 (33)	Lipiodol Ultra Fluid	>20ml of contrast, poor definition of fluoroscopy images	Pulmonary and retina	3 months of impaired vision
Dan et al., 1990 (41)	Lipiodol Ultra Fluid	Use of fluoroscopy not reported	Pulmonary, central nervous system	Comatose for 11 days, afterwards normal mental/motor function
Uzun et al., 2004 (11)	Lipiodol (not specified)	Use of fluoroscopy not reported	Pulmonary, central nervous system	Comatose for 10 days, afterwards mental/motor function progressively improved

HSG = hysterosalpingography

When including only the studies (including the case reports) that used fluoroscopy screening, there were 250 women with intravasation after an HSG with the use of oil-based contrast. In this group there were 16 women with oil embolisms (16/250, 6.4%), of which two had symptoms of coughing and one temporary impaired vision as a result of a retinal oil embolism (3/16, 18.8%). The authors reported that the fluoroscopy images were of poor quality, and over 20 ml of contrast was used during this last procedure (30).

When excluding the studies with known fluoroscopy guidance, there were 611 women with intravasation after an HSG with the use of oil-based contrast. In this group there were 24 women with oil embolisms (24/611, 3.9%), of which 19 (19/24, 79.2%) had, mostly transient, pulmonary symptoms. Of the 24 women with oil embolisms there

were three women with serious lasting consequences of cerebral complaints after an oil embolism (Table 1) (11, 33, 41).

Infection

Two RCT and 18 cohort studies reported on the frequency of infection after HSG with the use of oil-based contrast. The overall pooled weighted frequency of infection was 0.90% (95% CI 0.47–1.50, 70/11,287 women). Two RCT and two cohort studies compared HSG with the use of oil-based contrast to HSG using water-based contrast. The frequency of infection with the use of water-based contrast was 1.9% (95% CI 0.27–4.60, 17/564 women). Including only the studies published in or after 1960, the overall pooled frequency of infection was 0.55% (95% CI 0.23–1.00) after HSG with the use of oil-based contrast and 0.35% (95% CI 0.00–7.30) with the use of water-based contrast. The use of antibiotic prophylaxis was not systematically reported.

Mortality

Five cases of mortality were reported after HSG with the use of oil-based contrast in subfertile women. Four of these cases were infection-related, and they were published in the period between 1942 and 1950 (19, 40, 91). The fifth case described a woman that passed away minutes after a recurrent HSG with 9 ml of lipoiodine under light cyclopropane anaesthesia, possibly due to an allergic reaction to the oil-based contrast or the anaesthesia used (39). Additionally, two cases were reported in 1928 and 1930 where tubal blockage was found on the HSG. These women underwent surgery 1 and 5 days later, and died shortly after, presumably from infectious complications of the surgery (52, 79).

Lipogranuloma and oil remnants

Eleven studies reported on 41 women with lipogranuloma formation after an HSG with the use of different types of oil-based contrast. These included three cohort studies, one case series and seven case reports. The contrasts used were: Lipiodol not further specified (33 cases), oil-based/iodized contrast not further specified (five cases), Jodipin (two cases), Ethiodol (one case). In nine cases histology examination was mentioned, in 32 cases this was not mentioned.

Additionally, there were 85 reports of oil remnants after an HSG with the use of oil-based contrast. These were reported in nine studies; three cohort studies and six case reports. Forty-four cases were discovered within 2 weeks after the procedure, while 41 were discovered up to 27 years after the HSG procedure. Fifty-six cases were diagnosed after laparoscopy; 29 cases were diagnosed after radiology imaging. Histological examination was only reported in one case.

Thyroid dysfunction

Table 2 shows four cohort studies and four case reports/series on maternal thyroid function after HSG with the use of oil-based contrast.

Three cases of fetal goitre following an HSG with oil-based contrast were reported. In two of the cases the HSG had been performed in the month of conception (10 ml of Lipiodol and an unknown volume of unspecified oil-based contrast was used); in the third case three HSGs had been performed in the year before conception. In one case intra-amniotic levothyroxine was administered as treatment. After birth, hypothyroidism was diagnosed in one of the newborns, which resolved by day 7. The other neonates were euthyroid. One of the mothers had hypothyroidism during pregnancy; two were euthyroid. In one of the mothers, oil remnants were present in the abdominal cavity on a post-partum X-ray (81, 93, 116).

One retrospective cohort study (94) from Japan evaluated the neonatal thyroid function after HSG with the use of Lipiodol. Abnormal congenital thyroid screening was seen in 2.4% (5/212); three cases of subclinical hypothyroidism and two cases of overt hypothyroidism. The median volume of contrast in the group with thyroid dysfunction was significantly higher than the group with normal thyroid function (20 ml [range 10–20 ml] versus 8 ml [range 3–25 ml], P = 0.033). However, the volume was only reported for three out of five neonates with abnormal thyroid function test results.

Another retrospective cohort study investigated the thyroid function of 140 neonates born after a preconceptional HSG with oil-based contrast, Lipiodol Ultra Fluid (n = 76) or water-based contrast, Telebrix Hystero® (n = 64). None of the neonates tested positive during the congenital hypothyroidism screening. Furthermore, the volume of contrast used did not influence the thyroid function (median of 9.0 ml of oil-based contrast) (108).

Table 2. Maternal thyroid function after HSG

Study design		Procedure	Thyroid function pre-HSG	Outcome
Case reports				
Li et al., 2018 (64)	China	Oil-based contrast	Unknown	Fourteen women with increased urinary iodine content: 50% (7/14) (subclinical) hypothyroidism. All neonates tested negative during congenital thyroid screening
Sasaki et al., 2017 (93)	Japan	Oil-based contrast	Unknown	Case of hypothyroidism, no treatment. Fetal goiter.
Ma et al., 2016 (68)	China	Oil-based contrast 100ml	Euthyroid	Hyperthyroidism, no treatment, resolved spontaneously after 1.5 months
Ishizuki et al., 1992 (54)	Japan	Lipiodol	Graves' disease	Thyroiditis, goitre, treated with steroids for 2 months.
Cohorts/RCT				
So et al., 2017 (101)	Japan	Lipiodol Max 5ml	Euthyroid	Oil-based contrast: 22.6% subclinical hypothyroidism after 1–30 days, 24.4% after 31–180 days. Water-based contrast: 9.5% subclinical hypothyroidism after 1–30 days, 3.6% after 31–180 days.
Kaneshige et al., 2015 (56)	Japan	Lipiodol 6.1ml (4.0-9.0)	Euthyroid, 27% goiter palpable	0% hypothyroidism (0/22). 13.6% (3/22) transient subclinical hypothyroidism.
Mekaru et al., 2008 (73)	Japan	Lipiodol 5-10ml	76% euthyroid 12% subclinical hypothyroidism 12% subclinical hyperthyroidism	Euthyroid: 4/180 (2.2%) hypothyroidism, 28/180 (15.6%) subclinical hypothyroidism, 2/180 (1.1%) subclinical hyperthyroidism. Subclinical hypothyroidism: 10/28 (35.7%) hypothyroidism, three required thyroid hormone replacement. 1/28 (3.6%) subclinical hyperthyroidism. Subclinical hyperthyroidism: 4/12 (33.3%) normalization, 2/12 (16.7%) unchanged.
Slater et al., 1959 (100)	USA	Lipiodol	Clinically euthyroid	Oil-based contrast: 80% depression of iodine uptake, increase in protein-bound iodine for 4 months. Water-based contrast: no depression of iodine uptake. Increase in protein-bound iodine for 24–48 h.

HSG = hysterosalpingography

Other complications

One case of a tubal rupture, without ill effects, was described. The diagnostic method was not reported (114). Additionally, one case report described abdominal pain, like Fitz- Hugh–Curtis syndrome, possibly due to the chemical stimulation of the iodized oil (not further specified) used during an HSG (75).

HSG performed for non-subfertility indications

The primary intention of this study was to take into account HSGs performed for subfertility. However, in a non-systematic way, the study also identified one case of a massive oil embolism leading to death, published in 1931. A 60-year-old received an HSG with 8 ml Jodipin for postmenopausal blood loss which was suspected for malignancy. A massive oil embolism occurred in the cerebrum, pituitary gland, liver, spleen, kidney and heart, and the patient died within 5 h after the procedure. The use of fluoroscopy screening was not reported. It is likely that no adequate fluoroscopy was performed at the time (122).

Furthermore, a case report of a woman falling into a comatose state after an HSG was reported. This woman had had two unsuccessful curettage attempts for termination of pregnancy, after which she received an HSG with Lipiodol Ultra Fluid. The endometrium was injured after the several curettages, and so the contrast could flow directly into the bloodstream, leading to a massive intravasation with oil embolisms. After 81 days she was discharged with slight mental deficit (123).

This study also identified case reports of pulmonary oil embolisms after HSG performed in patients with: tubal ligation (n = 2) (124) , suspected endometrium carcinoma (n = 1) (125), abdominal pain (n = 1) (126), uterus myomatosus (n = 2) (127) and missed abortion (n = 1) (128).

DISCUSSION

In this review of articles published from 1928 onwards, including a total of 23,536 HSG with the use of oil-based contrast, the most frequently reported complication of HSG performed for subfertility was intravasation of contrast. This occurred in 2.7% of the HSG with the use of oil-based contrast (31 studies, 95% CI 1.7–3.8), compared with 2.0% with the use of water-based contrast (8 studies, 95% CI 1.2–3.0) derived from cohort studies and RCT. Oil embolisms occurred in 0.1% of the HSG performed in cohort studies and RCT. In all studies, including the case reports, the percentage of symptomatic oil embolisms was strikingly lower in the group with fluoroscopy guidance during HSG compared with

no fluoroscopy guidance (19% versus 79%). With the use of fluoroscopy guidance during HSG, no serious consequences of oil embolisms occurred.

The frequency of infection with the use of oil-based contrast was 0.90% (20 studies, 95% CI 0.47–1.50), compared with 1.9% (four studies, 95% CI 0.27–4.60) with the use of water-based contrast. One case of non-infection-related mortality after an HSG, most likely due to an anaphylactic reaction, was reported in 1947. There were 85 reports of oil remnants after an HSG. Half of the cases were diagnosed within 2 weeks of the procedure. Furthermore, there were 41 reports of lipogranuloma formation.

Women with subclinical hypothyroidism seem more likely to develop hypothyroidism after an HSG with oil-based contrast (35.7% versus 0–2.2% in euthyroid women), however this is based on only 28 and 202 women, respectively (56, 73, 101). Results on the effect on thyroid function of the offspring are contradictory; a Japanese study showed abnormal congenital thyroid screening in 2.4% whereas a Dutch study did not show any abnormalities (94, 108).

This is the first systematic review on the safety of HSG with oil-based contrast that includes all study types. Another strength of this systematic review is that no restriction on language or publication date was applied.

However, the systematic review has limitations. First, the quality of the included studies was moderate to low. This is attributable to the design and the publication year of the included studies. In most of the studies the primary outcome was pregnancy-related. Complications were often reported as secondary outcomes. Second, the development of fluoroscopy guidance during HSG has helped clinicians to diagnose intravasation and oil embolisms, leading to timely termination of the HSG procedure. This development is suggested as the reason for the increase in reported cases of intravasation and oil embolisms, however as mentioned previously, the percentage of symptomatic oil embolisms has therefore drastically decreased.

Oil embolisms, also known as fat embolisms, have not only been reported in the gynaecological literature. Bone marrow fat embolisms occur in 11–19% of trauma or orthopaedic surgery patients (129). Fat embolisms may cause a fat embolism syndrome, with clinical symptoms varying from right heart failure and cardiovascular collapse to hypoxemia, pyrexia, petechial rash and neurological symptoms (129) . When reaching the lungs, the fatty substance mixes with the locally secreted lipase. Free fatty acids are released, causing inflammation to the pulmonary microvasculature and leading to a shock lung-like or acute respiratory distress syndrome-like syndrome (130). Suggested

treatment is mainly supportive. Corticosteroids are proposed for their possible beneficial effect on the pulmonary capillary membrane, preventing pulmonary oedema (129). The pathogenesis of oil embolisms after the use of oil-based contrast could be similar to that described after a bone marrow fat embolism, however, in the latter case it concerns autologous tissue, while in the case of the use of oil-based contrast it concerns foreign material. In the four cases with severe complications of oil embolisms that are summarized in this review, one case was treated with corticosteroids (30), but in the other cases only supportive measures were reported.

In this systematic review of HSGs with oil-based contrast for subfertility, four cases of infection-related mortality were identified. It should be noted that these cases were all in the 1940s, when penicillin had been recently introduced and the treatment for infection was completely different from current practice (131). There are also reports in the literature of infection-related mortality following HSG with the use of water-based contrast (132). With the increased use and improvements of (prophylactic and therapeutic) antibiotics, the course of these infections has become less severe. The frequency of acute pelvic inflammatory disease after HSG is nowadays 0.5% with antibiotic prophylaxis and 1.4% without prophylaxis (133).

There were more than twice the number of reports on oil remnants (n = 85) than lipogranuloma formation (n = 41) after HSG with the use of oil-based contrast. Lipogranuloma is a pathological diagnosis and may be missed if oil remnants are not sent for pathological examination. Lipogranuloma may result in adhesion formation (47).

After iodine exposure, there is an excess of iodine transportation into the thyroid gland. Through negative feedback, this causes a transient decrease in the synthesis of thyroid hormone, potentially leading to the development of subclinical hypothyroidism. The level of thyroid hormone production will normally be restored within 24–48 h. However, patients with underlying thyroid abnormalities may be unable to escape from this so-called acute Wolff–Chaikoff effect and therefore acquire an iodine-induced (transient) overt hypothyroidism (13). This is in line with the results of the cohort study by Mekaru et al. (2008), which showed that 35.7% of women with a subclinical hypothyroidism develop overt hypothyroidism after an HSG with oil-based contrast, compared with 0–2.2% of euthyroid women (56, 73). Iodine induced (transient) hyperthyroidism can also occur in susceptible patients due to activation of quiescent nodules (13). This was shown in a case report of a woman with Graves' disease, who developed hyperthyroidism after an HSG (54).

Five out of eight studies included in this review, on maternal thyroid dysfunction after HSG, were performed in Japan. The effect of iodinated contrast on the thyroidal gland

may vary between Japanese and Caucasian women, possibly because of a different background risk (i.e. iodine-rich diet). The consumption of iodine-rich foods by mothers in Japan has been shown to lead to neonatal hypothyroidism (134). This may be reflected in the overall risk for congenital hypothyroidism, which is 0.7% in Japan compared with 0.04% in the Netherlands (135, 136). Data on Asian women suggest that neonatal thyroid dysfunction after HSG is related to the amount of oil-based contrast used during the procedure, although volume of contrast was not reported for all procedures (94).

It is unclear whether Caucasian women with an underlying thyroid disease are also at risk of developing hypothyroidism after an HSG with oil-based contrast. Until further studies have been performed, it is suggested that women with overt thyroid disease should not receive an HSG with oil-based contrast. In current practice, routine thyroid screening for women with subfertility varies. According to the NICE guidelines thyroid screening is not recommended as routine measurement in asymptomatic women presenting with subfertility (1). However, the ACOG committee opinion on fertility work-up does recommend routine thyroid testing for all subfertile women (137). Moreover, the 2017 American Thyroid Association guidelines for the diagnosis and management of thyroid disease during pregnancy and the post-partum period advises maintaining serum TSH concentrations below 2.5 mIU/l preconceptually in the subfertility setting (138).

In this systematic review of complications of HSG from 1928 onwards, the most frequently reported complication with oil-based contrast is intravasation, occurring in 2.7%. Only four cases of serious consequences of oil embolisms in subfertile women have been published since 1928. Therefore, safety concerns should not be the reason to deny the use of oil-based contrast for tubal testing in women with unexplained subfertility.

Further studies on the effect of oil-based contrast on maternal and neonatal thyroid function in Caucasian women are suggested. Furthermore, future research should investigate the mechanism of the pregnancy-enhancing effect of oil-based contrast. By gaining knowledge on the mechanism of action, it would be possible to determine which women would benefit most from an HSG with the use of oil-based contrast.

ACKNOWLEDGEMENTS

We would like to thank Mr B de Vries, Clinical Librarian at the Máxima MC for his assistance in developing the search strategies and his perseverance in retrieving the old manuscripts and Mrs J Dieleman for her statistical assistance. We want to thank our colleagues who helped to translate the non-English articles: Mrs G Bach, Professor JH Barker, Mrs IA Fomichev, Mrs L Jongmans, Dr C Nagata, Dr I Nedelcu, Dr MM Porath, Dr A Romano and Dr R Wang.

This work was an investigator-initiated study and partly funded by Guerbet, France. Guerbet is the manufacturer of Lipiodol® Ultra Fluid. The funders had no influence in the study design, data collection, the analyses performed or the interpretation of the study data.

REFERENCES

1. National Institute Care Excellence (NICE), 2017. Fertility problems:assessment and treatment. NICE Guidel. CG156.
2. Rindfleisch, W., 1910. Darstellung des cavum uteri. Berliner Klin. Wochenschrift 47, 780–781.
3. Cary, W.H., 1914. Note on determination of patency of fallopian tubes by the use of collargol and x-ray shadow. Am J Obs. Dis Woman Child 69, 462–464.
4. Nielsen, 1946. Injuries caused by hysterosalpingography. Acta Obstet. Gynecol. Scand. 26, 565–597.
5. Soules, M.R., Spadoni, L.R., 1982. Oil versus aqueous media for hysterosalpingography: A continuing debate based on many opinions and few facts. Fertil. Steril. 38, 1–11. https://doi.org/S0015-0282(16)46388-0 [pii]
6. Bonnemain, B., Guerbet, M., 1995. The history of Lipiodol (1901-1994) or How a medication may evolve with the times. Rev. Hist. Pharm. (Paris). 42, 159–170.
7. Simescu, M., Varciu, M., Nicolaescu, E., Gnat, D., Podoba, J., Mihaescu, M., Delange, F., 2002. Iodized oil as a complement to iodized salt in schoolchildren in endemic goiter in Romania. Horm. Res. 58, 78–82. https://doi.org/hre58078 [pii]
8. Fang, F., Bai, Y., Zhang, Y., Faramand, A., 2018. Oil-based versus water-based contrast for hysterosalpingography in infertile women: a systematic review and meta-analysis of randomized controlled trials. Fertil. Steril. 110, 15-160.e3. https://doi.org/10.1016/j.fertnstert.2018.03.021
9. Wang, R., van Welie, N., van Rijswijk, J., Johnson, N.P., Norman, R.J., Dreyer, K., Mijatovic, V., Mol, B.W., 2019. The effectiveness of tubal flushing with different contrast media on fertility outcomes: a systematic review and network meta-analysis. Ultrasound Obstet Gynecol 54, 172-181. https://doi.org/10.1002/uog.20238
10. Pujol y Brull, A., Vanrell, J., Carulla Riera, V., 1929. L'injection accidentelle du systeme veineux utero-ovarien au cours de l'hysterographie sur le vivant. J. Radiol. d'electrologie 13, 38–44.
11. Uzun, O., Findik, S., Danaci, M., Katar, D., Erkan, L., 2004. Pulmonary and cerebral oil embolism after hysterosalpingography with oil soluble contrast medium. Respirology 9, 134–136. https://doi.org///dx.doi.org/10.1111/j.1440-1843.2003.00524.x
12. Roest, I., van Welie, N., Mijatovic, V., Dreyer, K., Bongers, M., Koks, C., Mol, B.W., 2020. Complications after hysterosalpingography with oil- or water-based contrast: results of a nationwide survey. Hum. Reprod. Open. https://doi.org/10.1093/hropen/hoz045
13. Wolff, J., Chaikoff, I.L., 1948. Plasma inorganic iodide as a homeostatic regulator of thyroid function. J. Biol. Chem. 174, 555–564.
14. van den Boogaard, E., Vissenberg, R., Land, J.A., van Wely, M., van der Post, J.A., Goddijn, M., Bisschop, P.H., 2011. Significance of (sub)clinical thyroid dysfunction and thyroid autoimmunity before conception and in early pregnancy: a systematic review. Hum. Reprod. Update 17, 605–619. https://doi.org/10.1093/humupd/dmr024 [doi]
15. Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Altman, D., Antes, G., Atkins, D., Barbour, V., Barrowman, N., Berlin, J.A., Clark, J., Clarke, M., Cook, D., D'Amico, R., Deeks, J.J., Devereaux, P.J., Dickersin, K., Egger, M., Ernst, E., Gøtzsche, P.C., Grimshaw, J., Guyatt, G., Higgins, J., Ioannidis, J.P.A., Kleijnen, J., Lang, T., Magrini, N., McNamee, D., Moja, L., Mulrow, C., Napoli, M., Oxman, A., Pham, B., Rennie, D., Sampson, M., Schulz, K.F., Shekelle, P.G., Tovey, D., Tugwell, P., 2009. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med. https://doi.org/10.1371/journal.pmed.1000097
16. Hoy, D., Brooks, P., Woolf, A., Blyth, F., March, L., Bain, C., Baker, P., Smith, E., Buchbinder, R., 2012. Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement. J. Clin. Epidemiol. 65, 934–939. https://doi.org/10.1016/j.jclinepi.2011.11.014

17. Alper, M.M., Garner, P.R., Spence, J.E.H., Quarrington, A.M., 1986. Pregnancy rates after hysterosalpingography with oil- and water-soluble contrast media. Obstet. Gynecol. 68, 6–9.
18. Aznar, R.R., Rojas, A., Zarco, D.D., Guerrero, N.R., 1969. Intravasation as a complication in hysterosalpingography. Ginecol. Obstet. Mex. 26, 379–387.
19. Bang, J., 1950. Complications of hysterosalpingography 29, 383–399.
20. Barqawi, R., Bani-Irshaid, I., Bdor, A.-N., 2007. The effects of contrast media in patients undergoing salpingography on pregnancy rates. J. Bahrain Med. Soc. 19, 133–136.
21. Bateman, B.G., Jr., W.C.N., Kitchin, I.I.I.J., 1980. Intravasation during hysterosalpingography using oil-base contrast media. Fertil. Steril. 34, 439–443.
22. Bergin, J.H.E., 1951. The advantages and disadvantages of salpingography with particular reference to the use of diodone viscous. Br. J. Radiol. 24, 93–102.
23. Bersi, S., 1977. Granulomatous salpingitis following hysterosalpingography with "Lipiodol." Pathologica 69, 307–313.
24. Binder, I., Olanescu, A., Gusti, B.S., 1976. Pulmonary embolism after hysterosalpingography with lipiodol (Rumanian). Rev. Chir. Oncol. Radiol. O.R.L.Oftalmologie Stomatol. - Ser. Radiol. 15, 45–49.
25. Bohm, I., 2010. Contrast medium-induced leukopenia. Gynecol. Obstet. Invest. 70, 192. https://doi.org///dx.doi.org/10.1159/000316269
26. Böttger, H., Fleck, A., 1955. Foreign body reactions after hysterosalpingography with lipoid-soluble and water-soluble contrast media for roentgenography. Zentralbl. Gynakol. 77, 1172–1178.
27. Brent, K., Hadden, W.E., Weston-Webb, M., Johnson, N.P., 2006. After the FLUSH trial: a prospective observational study of lipiodol flushing as an innovative treatment for unexplained and endometriosis-related infertility. Aust. N. Z. J. Obstet. Gynaecol. 46, 293–297. https://doi.org/AJO595 [pii]
28. Brown, W.E., Jennings, A.F., Bradbury, J.T., 1949. The absorption of radiopaque substances used in hysterosalpingography. Am. J. Obstet. Gynecol. 58, 1041–1053.
29. Buytaert, P., Meulyzer, P., 1977. Is there still a place for hysterosalpingography in the exploration of infertility? J. Belge Radiol. 60, 339–343.
30. Charawanamuttu, A.M., Hughes-Nurse, J., Hamlett, J.D., 1973. Retinal embolism after hysterosalpingography. Br. J. Ophthalmol. 57, 166–169.
31. Claus, E., Dochez, C., 1966. Pulmonary lipiodol miliaria. Complications of hysterosalpingography and of vascular accidents. J. Belge Radiol. 49, 18–24.
32. Coventry, W.A., 1934. Accidental injection of the uteroovarian venous system during lipiodol uterosalpingography. Am. J. Obstet. Gynecol. 27, 912–913.
33. Dan, U., Oelsner, G., Gruberg, L., Ezra, D., Menczer, J., 1990. Cerebral embolization and coma after hysterosalpingography with oil-soluble contrast medium. Fertil. Steril. 53, 939–940.
34. Dreyer, K., Van, R.J., Mijatovic, V., Goddijn, M., Verhoeve, H.R., Van, R.I., Hoek, A., Bourdrez, P., Nap, A.W., Rijnsaardt-Lukassen, H.G.M., Timmerman, C.C.M., Kaplan, M., Hooker, A.B., Gijsen, A.P., Van, G.R., Van, H.C., Sluijmer, A. V, De, Bruin J, Smeenk, J.M.J., De, Boer J, Scheenjes, E., Duijn, A.E.J., Mozes, A., Pelinck, M.J., Traas, M.A.F., Van, H.M., Van, U.G., De, K.C., Van, G.N., Twisk, J.W.R., Hompes, P.G.A., Mol, B.W.J., 2017. Oil-based or water-based contrast for hysterosalpingography in infertile women. N. Engl. J. Med. 376, 2043–2052. https://doi.org///dx.doi.org/10.1056/NEJMoa1612337
35. Drukman, A., Rozin, S., 1951. Uterovenous and uterolymphatic intravasation in hysterosalpingography. J. Obstet. Gynaecol. (Lahore). 58, 73–78.
36. Effkemann, G., 1935. Uterovenöser Reflux bei Hysterosalpingographie. Arch. fuer Gynaekologie 160, 586–588.

37. Eisen, D., Goldstein, J., 1945. Lipiodol intravasation during uterosalpingography with pulmonary complications. Radiology 45, 603–607. https://doi.org/10.1148/45.6.603 [doi]
38. Elliott, G.B., Brody, H., Elliott, K.A., 1965. Implications of "lipoid salpingitis". Fertil. Steril. 16, 541–548. https://doi.org/10.1016/S0015-0282(16)35653-9
39. Faris, A.M., McMurrey, A., 1947. Uterosalpingography report of a fatality. Texas J. Med. 42, 592–597.
40. Feiner, D., 1942. A comparative study of tubal insufflation and Lipiodol injection in sterility. Am. J. Obstet. Gynecol. 43, 639–652.
41. Flew, J.D., 1944. Section of Obstetrics and Gynæcology. Proc. R. Soc. Med. 37, 425.
42. Fochem, K., Ulm, R., 1954. Studies on the tissue reaction in hysterosalpingographies with water-soluble contrast media. Fortschr. Geb. Rontgenstr. 80, 635–638.
43. Frischkorn, R., 1958. Embolism danger in use of oily contrast media with a study on the problem of contrast media in hysterosalpingography. Schweiz. Med. Wochenschr. 88, 1267–1269.
44. Geary, W.L., Holland, J.B., Weed, J.C., Jr., J.C.W., 1969. Uterosalpingography. Am. J. Obstet. Gynecol. 104, 687–692.
45. Gotoh, M., Yamaguchi, S., Hyodou, H., Oosuga, J., Kamei, Y., Hujij, T., 2010. A case of long-term residual oily contrast agent like as intra-abdominal foreign bodies. Kanto Soc. Obstet. Gynecol. 47, 366–366.
46. Grant, I.W., Callam, W.D., Davidson, J.K., 1957. Pulmonary oil embolism following hysterosalpingography. J. Fac. Radiol. Radiol. (Great Britain) 8, 410–415.
47. Grosskinsky, C.M., Clark, R.L., Wilson, P.A., Novotny, D.B., 1994. Pelvic granulomata mimicking endometriosis following the administration of oil-based contrast media for hysterosalpingography. Obstet. Gynecol. 83, 890–892.
48. Grossmann, M.., 1946. Pulmonary oil embolism. Br. J. Radiol. 19, 178–180.
49. Gunsberger, F., 1958. Case of embolism caused by iodized oil in hysterosalpingography. Med. Arh. 12, 49–52.
50. Heinen, G., Schussler, R., 1966. On the problem of accidental blood vessel filling in hysterosalpingography. Geburtshilfe Frauenheilkd. 26, 1178–1183.
51. Hemmeler, G., 1938. Lipiodolembolie in Gehirn und lungen nach Hysterosalpingographie. Schweiz. Med. Wochenschr. 25, 717–719.
52. Hirst, J.C., 1928. The technic of intrauterine lipiodol injections in gynecologic diagnosis. Am. J. Obstet. Gynecol. 15, 797–799.
53. Hohlbein, R., 1965. Gefass-influx mit ol-embolie bei hysterosalpinoggraphie. Zentralbl. Gynakol. 87, 370–374.
54. Ishizuki, Y., Hirooka, Y., Murata, Y., Tanigawa, S., 1992. A case of chronic thyroiditis with transient painful thyroiditis occurring after the administration of lipiodol. Nippon Naibunpi Gakkai zasshi 68, 1089–1095.
55. Johnson, N.P., Farquhar, C.M., Hadden, W.E., Suckling, J., Yu, Y., Sadler, L., 2004. The FLUSH trial - Flushing with lipiodol for unexplained (and endometriosis-related) subfertility by hysterosalpingography: A randomized trial. Hum. Reprod. 19, 2043–2051. https://doi.org/10.1093/humrep/deh418
56. Kaneshige, T., Arata, N., Harada, S., Ohashi, T., Sato, S., Umehara, N., Saito, T., Saito, H., Murashima, A., Sago, H., 2015. Changes in serum iodine concentration, urinary iodine excretion and thyroid function after hysterosalpingography using an oil-soluble iodinated contrast medium (lipiodol). J. Clin. Endocrinol. Metab. 100, 469. https://doi.org/10.1210/jc.2014-2731 [doi]
57. Karshmer, N., Stein, W., 1951. Oil embolism during hysterosalpingography. Am. J. Obstet. Gynecol. 61, 458–460. https://doi.org/10.1016/0002-9378(51)90272-4

58. Kika, K., 1954. A clinical analysis of the “angiograms” found in the course of hysterosalpingography with special reference to tuberculosis of the female genitals. Am. J. Obstet. Gynecol. 67, 56–63.

59. Kilroe, J.C., Hellman, A.M., 1933. Entrance of lipiodol into ovarian and other veins during uterography. Am. J. Obstet. Gynecol. 25, 152–153.

60. Kuzavova, N.I., 1964. Intravasation during hysterosalpingography. Vestn. Rentgenol. Radiol. 39, 51–54.

61. La Sala, G., Ghirardini, G., Valli, F., Margini, F., 1982. Intravasation during hysterosalpingography using low viscosity oil base contrast media. Clin. Exp. Obstet. Gynecol. 9, 257–259.

62. Lau, H.U., 1969. Accidental demonstration of the vessels in hysterosalpingography. Zentralbl. Gynakol. 91, 273–277.

63. Levinson, J.M., 1963. Pulmonary Oil Embolism Following Hysterosalpingography
. Fertil. Steril. 14.

64. Li, R., Liu, Y., Ma, L., Qiu, L., Han, J., 2018. Excessive exposure to iodine in pregnancy merits attention: A pilot follow-up study. Int. J. Gynecol. Obstet. 143, 175–176.

65. Lin, Y.Y., Tsou, S.H., 1935. The escape of Lipiodol into the utero-ovarian venous system in hysterosalpingography: with special emphasis on the pulmonary complications. Chinese M. J. 49, 1241–1250.

66. Lindequist, S., Rasmussen, F., Larsen, C., 1994. Use of iotrolan versus ethiodized poppy-seed oil in hysterosalpingography. Radiology 191, 513–517.

67. Liu, L.-X., Zhao, J.-H., Zhang, G.-F., 2010. Comparison of iodized oil with soluble iodized agents in hysterosalpingography and the evaluation of modified hysterosalpingography. J. Interv. Radiol. 19, 574–577.

68. Ma, G., Mao, R., Zhai, H., 2016. Hyperthyroidism secondary to hysterosalpingography: an extremely rare complication: A case report. Medicine (Baltimore). 95, e5588. https://doi.org/10.1097/MD.0000000000005588 [doi]

69. Mackey, R.A., Glass, R.H., Olson, L.E., Vaidya, R., 1971. Pregnancy following hysterosalpingography with oil and water soluble dye. Fertil. Steril. 22, 504–507. https://doi.org/S0015-0282(16)38404-7 [pii]

70. Madsen, V., 1942. Kliniske studier over hyaterosalpingografi.

71. Malter, I.J., Fox, R.M., 1972. Prolonged oviduct retention of iodized contrast medium. Obstet. Gynecol. 40, 221–224.

72. Meaker, S.R., 1934. Accidental injection of iodized oil into uterine veins. Am. J. Obstet. Gynecol. 28, 568–571.

73. Mekaru, K., Kamiyama, S., Masamoto, H., Sakumoto, K., Aoki, Y., 2008. Thyroid function after hysterosalpingography using an oil-soluble iodinated contrast medium. Gynecol. Endocrinol. 24, 498–501. https://doi.org///dx.doi.org/10.1080/09513590802246364

74. Miyazaki, Y., Yamamoto, T., Hyakudomi, R., Taniura, T., Hirayama, T., Takai, K., Hirahara, N., Tajima, Y., 2020. Case of inflammatory granuloma in inguinal hernia sac after hysterosalpingography with oily contrast medium. Int. J. Surg. Case Rep. https://doi.org/10.1016/j.ijscr.2020.05.084

75. Morii, K., Yamamoto, T., Kishida, H., Okushin, H., 2013. Fitz-hugh-curtis syndrome-like findings: Chemical stimulation by the contrast agent. Intern. Med. 52, 2587–2588. https://doi.org///dx.doi.org/10.2169/internalmedicine.52.1173

76. Netter, A., Weill-Fage, J.C., 1950. Accidental injection of lipiodol into the paravertebral lymphatics, with a digression on certain lumbo-sacral pains in women. Bull. Mem. Soc. Med. Hop. Paris.

77. Nordio, A., 1938. Quadri radiologici rari di isterografia. Liquido opaco nei vasi uterini. Atti della Soc. Ital. di Ostet. e Ginecol. 34, 565–567.

78. Norris, S., 1956. The hysterogram in study of sterility. Canad 75, 1016–1020.
79. Novak, J., 1930. Salpingographie oder Tubendurchblasung? Zentralbl. Gynakol. 54, 3013–3017.
80. Nugent, D., Watson, A.J., Killick, S.R., Balen, A.H., Rutherford, A.J., 2002. A randomized controlled trial of tubal flushing with lipiodol for unexplained infertility. Fertil. Steril. 77, 173–175. https://doi.org///dx.doi.org/10.1016/S0015-0282%2801%2902925-9
81. Omoto, A., Kurimoto, C., Minagawa, M., Shozu, M., 2013. A case of fetal goiter that resolved spontaneously after birth. J. Clin. Endocrinol. Metab. 98, 3910–3911. https://doi.org///dx.doi.org/10.1210/jc.2013-1066
82. Palmer, A., 1960. Ethiodol hysterosalpingography for the treatment of infertility. Fertil. Steril. 11, 311–315. https://doi.org/10.1016/s0015-0282(16)33788-8
83. Pear, B.L., Boyden, F.M., 1967. Intraperitoneal lipid granuloma. Radiology 89, 47–51.
84. Piatt, A.D., 1947. Intravasation of lipiodol during uterosalpingography. Ohio State Med. J. 43, 821–824.
85. Porcher, M.P., 1935. Un incident de salpingographie. Bull. Mem. la Soc. d'Electroradiologie Medicale Fr. 23, 146–147.
86. Rasmussen, F., Justesen, P., Nielsen, D.T., 1987. Therapeutic value of hysterosalpingography with lipiodol ultra fluid. Acta radiol. 28, 319–322.
87. Riche, V., Fayot, G., 1931. Le lipiodiagnostic en gynecologie. Gynecol. Obstet. (Paris). 24, 540–542.
88. Ries, E., 1929. Effect of Lipiodol Injection on the Tubes. Am. J. Obstet. Gynecol. 17, 728–730.
89. Robins, S.., Shapira, A.., 1951. Uterotubography.
90. Rubin, I.C., 1928. Diagnostic use of intra-uterine iodized oil injection combined with the X-rays, as compared to peruterine CO2 insufflations. Radiology 11, 115–125.
91. Rutherford, R.N., 1948. The therapeutic value of repetitive Lipiodal tubal insufflations. West. J. surgery, Obstet. Gynecol. 56, 145–154.
92. Sappey, P., Fabre, H., Jouffrey, A., 1952. Accidental vasculary injection of lipiodol in the course of hysterosalpingography. Gynecol. Obstet. (Paris). 51, 298–300.
93. Sasaki, Y., Kikuchi, A., Murai, M., Kanasugi, T., Isurugi, C., Oyama, R., Sugiyama, T., 2017. Fetal goiter associated with preconception hysterosalpingography using an oil-soluble iodinated contrast medium. Ultrasound Obstet. Gynecol. 49, 275–276. https://doi.org/10.1002/uog.15902
94. Satoh, M., Aso, K., Katagiri, Y., 2015. Thyroid Dysfunction in Neonates Born to Mothers Who Have Undergone Hysterosalpingography Involving an Oil-Soluble Iodinated Contrast Medium. Horm. Res. Paediatr. 84, 370–375. https://doi.org/10.1159/000439381
95. Schaffer, B., 1954. Post-hysterosalpingographical giant-cell granuloma of the fallopian tubes]. Obstet. Ginecol. Lat. Am. 12, 562–565.
96. Schultze, G.K.F., 1932. Schädigungen und Gahren bei der Hysterosalpingographie. Zeitschrift für Geburtshilfe und Gynaekologie 101, 413–436.
97. Schutte, J.M., Oonk, M.H., van der Ploeg, J.M., 2006. Diagnosis image (258). A women with a "bullet" in the abdomen. Ned. Tijdschr. Geneeskd. 150, 143.
98. Schwabe, M.G., Shapiro, S.S., Haning, R. V, 1983. Hysterosalpingography with oil contrast medium enhances fertility in patients with infertility of unknown etiology. Fertil. Steril. 40, 604–606.
99. Shapiro, J.H., Rubinstein, B., Jacobson, H.G., Popei, H.M., 1957. Pulmonary oil embolism: a complication of hysterosalpingography. Am. J. roentgenoloy, radium Ther. Nucl. Med. 77, 1055–1058.
100. Slater, S., Paz-Carranza, J., Solomons, E., Perlmutter, M., 1959. Effect of hysterosalpingography on assay of thyroid function. Fertil. Steril. 10, 144–149.

101. So, S., Yamaguchi, W., Tajima, H., Nakayama, T., Tamura, N., Kanayama, N., Tawara, F., 2017. The effect of oil and water-soluble contrast medium in hysterosalpingography on thyroid function. Gynecol. Endocrinol. 33, 682–685. https://doi.org/10.1080/09513590.2017.1307960 [doi]

102. Solal, R., 1932. Injection accidentelle de la veine iliaque au cours d'une hysterosalpingographie. Soc. d'obstetrique Gynecol. 21, 725–731.

103. Steiner, A.Z., Meyer, W.R., Clark, R.L., Hartmann, K.E., 2003. Oil-soluble contrast during hysterosalpingography in women with proven tubal patency. Obstet. Gynecol. 101, 109–113. https://doi.org///dx.doi.org/10.1016/S0029-7844%2802%2902390-6

104. Stoll, P., Zeitz, H., 1956. Injuries caused by contrast media after hysterosalpingography with iodipin; histologic examination of 33 excised tubes. Dtsch. Med. Wochenschr. 81, 1557–1560.

105. Takeyama, K., Ishikawa, R., Nakayama, K., Suzuki, T., 2014. Intraperitoneal residual contrast agent from hysterosalpingography detected following cesarean section. Tokai J. Exp. Clin. Med. 39, 69–71.

106. Tan, Y., Zheng, S., Lei, W., Wang, F., Jiang, S., Zeng, T., Zhou, B., Hong, F., 2019. Ethiodized poppyseed oil versus ioversol for image quality and adverse events in hysterosalpingography: A prospective cohort study, in: BMC Medical Imaging. https://doi.org/10.1186/s12880-019-0346-0

107. Ueda, M., Koshiyama, M., Sato, Y., 2016. Prophylactic corticosteroids in pulmonary oil embolism after hysterosalpingography. J. Obstet. Gynaecol. (Lahore). 36, 137–138. https://doi.org/10.3109/01443615.2015.1041892

108. van Welie, N., Roest, I., Portela, M., van Rijswijk, J., Koks, C., Lambalk, C.B., Dreyer, K., Mol, B.W.J., Finken, M.J.J., Mijatovic, V., 2020. Thyroid function in neonates conceived after hysterosalpingography with iodinated contrast. Hum. Reprod. 35, 1159–1167. https://doi.org/10.1093/humrep/deaa049

109. Vara, P., 1950. Venograms found in connection with hysterosalpingography 39, 228–238.

110. Weise, W., Buttner, H.H., Echtermeyer, H., 1973. Fourteen-year-old contrast media deposits in both Fallopian tubes following hysterosalpingography. Zentralbl. Gynakol. 95, 457–462.

111. Weitzner, G., 1935. Entrance of iodized oil into the venous circulatory system during uterography. J. Am. Med. Assoc. 104, 545–546.

112. Werner, H., 1952. On the problem of harmfulness of the contrast oil in hysterosalpingography. Zentralbl. Gynakol. 74, 1952–1954.

113. Williams, E.., 1944. Venous intravasation during utero-salpingography. Br. J. Obstet. Gynaecol. 17, 13–17. https://doi.org/10.1148/53.3.410

114. Witwer, E.R., Cushman, H.P., Leucutia, T., 1930. The present status of hysterosalpingography. Am. J. Roentgenol. Radium Ther. 23, 125–159.

115. Wong, A.I.H., Wu, C., Chien, M.H., 1932. The escape of lipiodol into the utero-ovarian venous system in hystero-salpingography. Chin. Med. J. (Engl). 46, 162–167.

116. Yamazaki, T., Hyodo, M., Takelshi, N., 2019. A case of intra-amniotic levothyroxine treatment for fetal goiter after preonceptional hysterosalpingography using an oil-soluble iodinated contrast medium. J. Obstet. Gynaecol. Res. 45, 1703.

117. Zachariae, F., 1955. Venous and Lymphatic Intravasation in Hysterosalpingography.

118. Zacharin, D., 1933. Accidents in pyelography and hysterosalpingography: the venous escape of radiographic contrast media. Med. J. Aust. 2, 449–450.

119. Nunley, W.C., Bateman, B.G., Kitchin III, J.D., Pope, T.L., 1987. Intravasation during hysterosalpingography using oil-base contrast medium--a second look. Obstet. Gynecol. 70, 309–312.

120. Volk, W., 1936. Leistung und Gefahren der Hysterosalpinographie. Zeitschrift für Geburtshilfe und Gynaekologie 113, 339–372.

121. Woltz, J.H.., Bradford, W.Z., Bradford, W.B., McCoy, J.B., 1958. Complications of hysterosalpingography. Am. J. Obstet. Gynecol. 76, 736–741.

122. Gajzago, E., 1931. Ein im Anschluss an Hysterographie durch Ölembolie verursachter Todesfall. Zentralbl. Gynakol. 55, 543–544.

123. Ogihara, T., Miyao, H., Katoh, H., Ikenaga, H., Michikawa, N., Ohkubo, N., Kasahara, M., Iyori, S., 1991. Adverse reactions to lipiodol ultra fluid: Report of an accidental case. Keio J. Med. 40, 94–96.

124. Roblee, M.A., 1945. Lipiodol pulmonary emboli following hysterosalpingography. South. Med. J. 38, 89–94.

125. Breitländer, Hinrichs, 1941. Lungenembolie nach Hysterosalpingographie mittels Jodipin - reaktionsloser Verlauf. Zentralbl. Gynakol. 3, 124–129.

126. Ingersoll, F.M., Robbins, L.L., 1947. Oil embolism following hysterosalpingography. Am. J. Obstet. Gynecol. 53, 307–311.

127. Hodge, J., Price, A.H., 1969. The management of massive pulmonary embolism following hysterosalpingography. J. S. C. Med. Assoc. 65, 315–319.

128. Hinaut, G., Bonhomme, J., Merlat-Guitard, M., Mignon, M., Morel, B., 1966. Embolie graisseuse provoquee par une hysterographie. J. Fr. Med. Chir. Thorac. 21, 227–232.

129. Mellor, A., Soni, N., 2001. Fat embolism. Anaesthesia 56, 145–154.

130. Duran, H., Cardenas-Camarena, L., Bayter-Marin, J.E., Ramos-Gallardo, G., Robles-Cervantes, J.A., 2018. Microscopic and Macroscopic Fat Embolism: Solving the Puzzle with Case Reports. Plast. Reconstr. Surg. 142, 569e-577e. https://doi.org/10.1097/PRS.0000000000004810

131. Bud, R., 2007. Penicillin: triumph and tragedy, Oxford University Press. https://doi.org/10.1017/s0025727300003379

132. Lachmann, A., 1944. Et Dodsfald efter Hysterosalpingografi. Nord. Med. 22, 929–931.

133. Li, H.-M., Sung, F.-C., Li, S.-C., Huang, Y.-K., Chang, Y., Chang, C.-C., Huang, S.J., Lin, C.-L., Kao, C.-H., 2018. The effect of antibiotic prophylaxis for acute pelvic inflammatory disease after hysterosalpingography: a retrospective cohort study. Curr. Med. Res. Opin. 34, 1271–1276. https://doi.org/10.1080/03007995.2017.1417243

134. Nishiyama, S., Mikeda, T., Okada, T., Nakamura, K., Kotani, T., Hishinuma, A., 2004. Transient hypothyroidism or persistent hyperthyrotropinemia in neonates born to mothers with excessive iodine intake. Thyroid 14, 1077–1083. https://doi.org/10.1089/thy.2004.14.1077 [doi]

135. Activity Report Tokyo Health Service Association, 2010. 39, 133.

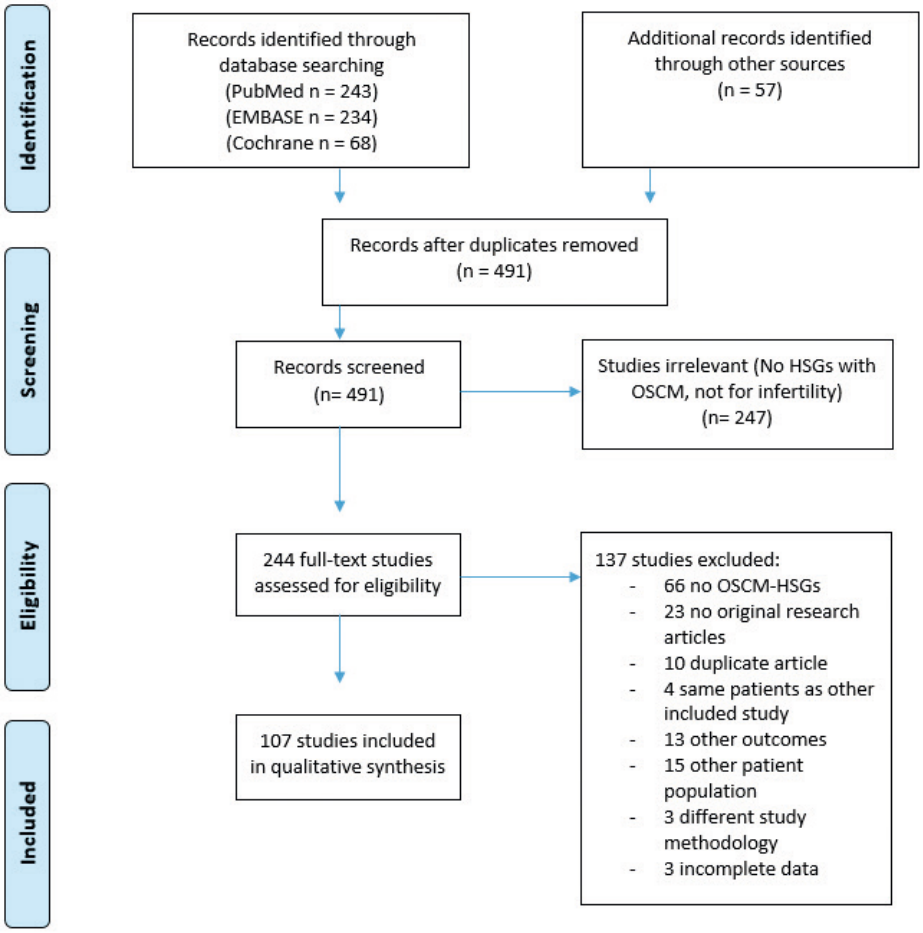
136. Verkerk, P.H., van Trotsenburg, A.S., Hoorweg-Nijman, J.J., Oostdijk, W., van Tijn, D.A., Kempers, M.J., van den Akker, E.L., Loeber, J.G., Elvers, L.H., Vulsma, T., 2014. Neonatal screening for congenital hypothyroidism: more than 30 years of experience in the Netherlands. Ned. Tijdschr. Geneeskd. 158, A6564.

137. ACOG, 2019. Infertility Workup for the Women's Health Specialist. ACOG committee opinion. Obstet. Gynecol. 133, 377–84. https://doi.org/10.1097/aog.0000000000003272

138. Alexander, E.K., Pearce, E.N., Brent, G.A., Brown, R.S., Chen, H., Dosiou, C., Grobman, W.A., Laurberg, P., Lazarus, J.H., Mandel, S.J., Peeters, R.P., Sullivan, S., 2017. 2017 Guidelines of the American Thyroid Association for the Diagnosis and Management of Thyroid Disease During Pregnancy and the Postpartum. Thyroid 27, 315–389. https://doi.org/10.1089/thy.2016.0457

SUPPLEMENTARY MATERIALS

Supplementary Figure 1. Flowchart of inclusions



Abbreviations: OSCM = oil-soluble contrast media.

Supplementary Table 1. Search PubMed 23-6-2020

Search	Add to builder	Query	Items found	Time
#10	Add	Search **#1 AND #4 AND #9**	243	04:02:22
#9	Add	Search **#7 OR #8**	5348646	03:50:34
#8	Add	Search **((adverse[tw] OR side[tw] OR undesirable[tw] OR injurious[tw]) AND (effect[tw] OR effects[tw])) OR risk[tw]**	4211290	03:49:14
#7	Add	Search **#5 OR #6**	1729799	03:48:50
#6	Add	Search **thyroid[tiab] OR hypothyroidism[tiab] OR intravasation[tiab] OR extravasation[tiab] OR embolization[tiab] OR embolisation[tiab] OR complication*[tiab] OR anaphylaxis[tiab] OR anaphylactic[tiab] OR pelvic inflammatory disease*[tiab] OR adnexitis[tiab] OR lipogranuloma[tiab]**	1239799	03:48:36
#5	Add	Search **"Thyroid Hormones"[Mesh] OR "Thyroid Diseases"[Mesh] OR "Thyroid Function Tests"[Mesh] OR "Hypothyroidism"[Mesh] OR "Extravasation of Diagnostic and Therapeutic Materials"[Mesh] OR "Pregnancy Complications"[Mesh] OR "Anaphylaxis"[Mesh] OR "Pelvic Inflammatory Disease"[Mesh] OR "Granuloma"[Mesh]**	691814	03:48:17
#4	Add	Search **#2 OR #3**	125755	03:47:56
#3	Add	Search **ethiodized oil*[tiab] OR lipiodol[tiab] OR ethiodol[tiab] OR iodinated contrast[tiab] OR oil contrast[tiab] OR oil soluble contrast[tiab] OR iodolipol[tiab] OR iodized oil*[tiab]**	7201	03:47:42
#2	Add	Search **"Ethiodized Oil"[Mesh] OR "Iodopyridones"[Mesh] OR "Iodized Oil"[Mesh] OR "Contrast Media" [Pharmacological Action] OR "Contrast Media"[Mesh]**	123400	03:47:25
#1	Add	Search **"Hysterosalpingography"[Mesh] OR hysterosalpingograph*[tiab] OR salpingograph*[tiab] OR HSG[tiab] OR uterosalpingograph*[tiab]**	5587	03:47:09

Supplementary Table 2. Search EMBASE (Ovid) 23-6-2020

# ▲	Searches	Results
1	exp hysterosalpingography/	4219
2	(hysterosalpingograph* or salpingograph* or HSG or uterosalpingograph*).ab,ti.	3864
3	1 or 2	5869
4	exp ethiodized oil/ or exp iodopyridone/ or exp iodinated poppyseed oil/ or exp contrast medium/	163382
5	(ethiodized oil* or lipiodol or ethiodol or iodinated contrast or oil contrast or oil soluble contrast or iodolipol or iodized oil*).ab,ti.	9576
6	4 or 5	165360
7	exp thyroid hormone/ or exp thyroid disease/ or exp thyroid function test/ or exp hypothyroidism/ or exp contrast medium extravasation/ or exp pregnancy complication/ or exp anaphylaxis/ or exp pelvic inflammatory disease/ or exp lipogranuloma/ or exp adnexitis/	425013
8	(thyroid or hypothyroidism or intravasation or extravasation or embolization or embolisation or complication* or anaphylaxis or anaphylactic or pelvic inflammatory disease* or adnexitis or lipogranuloma).ab,ti.	1633099
9	7 or 8	1829337
10	adverse.mp. or exp adverse event/ or exp adverse outcome/ or exp side effect/ or exp risk/	4309515
11	(((adverse or side or undesirable or injurious) and (effect or effects)) or risk).ab,kw,ti,sh.	3761777
12	10 or 11	5634843
13	9 or 12	6786477
14	3 and 6 and 13	234

Supplementary Table 3. Search Cochrane 23-6-2020

#1	MeSH descriptor: [Hysterosalpingography] explode all trees		MeSH	127
#2	(hysterosalpingograph* or salpingograph* or uterosalpingograph*):ti,ab,kw (Word variations have been searched)	S	Limits	385
#3	#1 OR #2		Limits	385
#4	MeSH descriptor: [Iodized Oil] explode all trees		MeSH	190
#5	MeSH descriptor: [Iodopyridones] explode all trees		MeSH	0
#6	MeSH descriptor: [Contrast Media] explode all trees		MeSH	2673
#7	(ethiodized oil* or lipiodol* or ethiodol or iodinated contrast or oil contrast):ti,ab,kw (Word variations have been searched)	S	Limits	1636
#8	(oil soluble contrast or iodolipol or iodized oil*):ti,ab,kw (Word variations have been searched)	S	Limits	216
#9	#4 or #5 or #6 or #7 or #8		Limits	3984
#10	#3 and #9		Limits	68

Supplementary Table 4. Qualiity assessment

Quality assessment checklist for prevalence studies (adapted from Hoy et al)

Name of author(s): Year of publication: Study title:		
Risk of bias items	**Risk of bias levels**	**Points scored**
1. Was the study's target population a close representation of the national population in relation to relevant variables, e.g. age, sex, occupation?	**Yes (LOW RISK)**: The study's target population was a close representation of the national population.	0
	No (HIGH RISK): The study's target population was clearly NOT representative of the national population.	1
2. Was the sampling frame a true or close representation of the target population?	**Yes (LOW RISK)**: The sampling frame was a true or close representation of the target population.	0
	No (HIGH RISK): The sampling frame was NOT a true or close representation of the target population.	1
3. Was some form of random selection used to select the sample, OR, was a census undertaken?	**Yes (LOW RISK)**: A census was undertaken, OR, some form of random selection was used to select the sample (e.g. simple random sampling, stratified random sampling, cluster sampling, systematic sampling).	0
	No (HIGH RISK): A census was NOT undertaken, AND some form of random selection was NOT used to select the sample.	1
4. Was the likelihood of non-response bias minimal?	**Yes (LOW RISK)**: The response rate for the study was ≥75%, OR, an analysis was performed that showed no significant difference in relevant demographic characteristics between responders and non- responders	0
	No (HIGH RISK): The response rate was <75%, and if any analysis comparing responders and non-responders was done, it showed a significant difference in relevant demographic characteristics between responders and non-responders	1
5. Were data collected directly from the subjects (as opposed to a proxy)?	**Yes (LOW RISK)**: All data were collected directly from the subjects.	0
	No (HIGH RISK): In some instances, data were collected from a proxy.	1
6. Was an acceptable case definition used in the study?	**Yes (LOW RISK)**: An acceptable case definition was used.	0
	No (HIGH RISK): An acceptable case definition was NOT used	1
7. Was the study instrument that measured the parameter of interest (e.g. prevalence of low back pain) shown to have reliability and validity (if necessary)?	**Yes (LOW RISK)**: The study instrument had been shown to have reliability and validity (if this was necessary), e.g. test-re- test, piloting, validation in a previous study, etc.	0
	No (HIGH RISK): The study instrument had NOT been shown to have reliability or validity (if this was necessary).	1
8. Was the same mode of data collection used for all subjects?	**Yes (LOW RISK)**: The same mode of data collection was used for all subjects.	0
	No (HIGH RISK): The same mode of data collection was NOT used for all subjects.	1
9. Were the numerator(s) and denominato r(s) for the parameter of interest appropriate	**Yes (LOW RISK)**: The paper presented appropriate numerator(s) AND denominator(s) for the parameter of interest (e.g. the prevalence of low back pain).	0
	No (HIGH RISK): The paper did present numerator(s) AND denominator(s) for the parameter of interest but one or more of these were inappropriate.	1
10. Summary on the overall risk of study bias	**LOW RISK**	0-3
	MODERATE RISK	4-6
	HIGH RISK	7-9

Supplementary Table 5. Demographics of included studies.

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Hirst, 1928)	**USA**	Case report	Case of death after surgery following-HSG	OSCM = 1	-	-	Lipiodol (not spec)	6	-	-
(Ries, 1929)	**USA**	Case report	Case of lipogranuloma	OSCM = 1	Subfertility (not spec)	30	Lipiodol (not spec)	-	-	Uni/bilateral patent tubes
(Pujol y Brull* et al., *1929)	**Spain**	Case series	Cases of intravasation	OSCM = 5	1 metrorrhagia, 1 possibly TBC	-	Lipiodol (not spec)	-	-	-
(Novak, 1930)	**Austria**	Case report	Case of death after surgery following-HSG	OSCM=1	Sec subfertility	36	Jodipin	-	-	Bilateral tubal occlusion
(Solal, 1932)	**France**	Case report	Case of intravasation	OSCM = 1	Subfertility (not spec)	30	Lipiodol (not spec)	8	-	Bilateral tubal occlusion Uterus infantile bicornus
(Wong* et al., *1932)	**China**	Case series	Cases of intravasation	OSCM = 4	50% prim subfertility 50% sterility not spec	22, 24, 31, 35	Lipiodol (not spec)	4, 8, 9, 15	No	Bilateral tubal occlusion 75% Unilateral tubal occlusion 25%
(Kilroe and Hellman, 1933)	**USA**	Case report	Case of intravasation	OSCM = 1	Prim subfertility	30	Lipiodol (not spec)	-	-	-

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Zacharin, 1933)	**Australia**	Case report	Case of intravasation	OSCM = 1	Subfertility (not spec) HSG after Rubin's test	30	Lipiodol (not spec)	8	-	Bicornuate uterus, unilateral tubal occlusion
(Coventry, 1934)	**USA**	Case report	Case of intravasation	OSCM = 1	Prim subfertility	34	Lipiodol (not spec)	10	-	Bilateral tubal occlusion
(Meaker, 1934)	**USA**	Case report	Case of intravasation	OSCM = 1	Prim subfertility	-	Ioidized oil (not spec)	11	-	Bilateral tubal occlusion
(Effkemann, 1935)	**Germany**	Case report	Case of intravasation	OSCM = 1	Subfertility (not spec)	34	Jodipin (20%)	12	-	Unilateral tubal occlusion
(Lin and Tsou, 1935)	**China**	Case Series	Cases of intravasation	OSCM = 7	Subfertility (not spec) 1 fistula after hysterotomy	31 (27-35)	Lipiodol (not spec)	8 (8-20)	No	Bilateral tubal occlusion 14% Patent 28% Not reported 57%
(Porcher, 1935)	**France**	Case report	Case of intravasation	OSCM = 1	Subfertility (not spec)	-	Lipiodol (not spec)	7.5	No	Bilateral patency
(Weitzner, 1935)	**USA**	Case report	Case of intravasation	OSCM = 1	Prim subfertility and myoma	35	Ioidized poppy-seed oil	8	-	1 patent, 1 previously removed
(Hemmeler, 1938)	**Switzerland**	Case report	Case of pulmonary embolism	OSCM = 1	Prim subfertility	30	Lipiodol (not spec)	4	-	Tubes not visible

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Flew, 1944)	**UK**	Case report	Case of pulmonary/ cerebral embolism	OSCM = 1	Subfertility (not spec)	-	Lipiodol (not spec)	-	-	-
(Williams, 1944)	**UK**	Case series	Cases of intravasation	OSCM = 6	Subfertility (not spec)	< 36	OSCM (not spec)	-	-	Bilateral tubal occlusion 17% Not reported 83%
(Eisen and Goldstein, 1945)	**Canada**	Case report	Case of pulmonary embolism	OSCM = 1	Sec subfertility	28	Lipiodol (not spec)	15	-	1 tube patent 1 previously removed
(Grossmann, 1946)	**Czech Republic**	Case series	Cases of intravasation	OSCM = 2	1 sec subfertility 1 subfertility (not spec)	34, unknown	Lipiodol (not spec) Neohydriol	-	-	1 patent tubes, 1 unknown
(Faris and McMurrey, 1947)	**USA**	Case series	Cases of intravasation / anaphylactic reaction	OSCM = 2	Prim subfertility 50% Unknown 50%	35, unknown	Lipiodine	10, 12	Yes	Bilateral tubal occlusion 50%. Not reported 50%
(Piatt, 1947)	**USA**	Case report	Case of intravasation	OSCM = 1	Sec subfertility	26	Lipiodol (not spec)	10	-	Unilateral patent tubes (1 tube previously removed)
(Netter and Weill-Fage, 1950)	**France**	Case report	Case of intravasation	OSCM = 1	Subfertility (not spec)	37	Lipiodol (not spec)	-	-	-

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Karshmer and Stein, 1951)	**UK**	Case report	Case of pulmonary emboli	OSCM = 1	Prim subfertility	29	Iodochlorol/ Lipiodol	11	No	Bilateral tubal occlusion
(Sappey* et al., *1952)	**France**	Case report	Case of pulmonary emboli	OSCM = 1	Subfertility (not spec)	-	Lipiodol 40%	-	-	-
(Werner, 1952)	**Germany**	Case report	Case of oil granuloma	OSCM = 1	-	24	Jodipin	-	-	-
(Fochem and Ulm, 1954)	**Germany**	Case series	Cases of intravasation	OSCM = 3	-	28, 36, 41	Jodipin	-	-	-
(Schaffer, 1954)	**Argentina**	Case report	Case of PID / foreign body granuloma	OSCM = 1	Sec subfertility	36	OSCM	-	-	Bilateral tubal occlusion + bilateral hydrosalpinx
(Böttger and Fleck, 1955)	**Germany**	Case report	Case of PID	OSCM = 1	Prim subfertility	-	OSCM	-	-	-
(Grant* et al., *1957)	**Scotland**	Case report	Case of pulmonary embolism	OSCM = 1	Prim subfertility	29	Lipiodol-Lafay	7-8	-	Both tubes patent
(Shapiro* et al., *1957)	**USA**	Case report	Case of intravasation	OSCM = 1	Sec subfertility and fibroids	26	Lipiodol (not spec)	-	-	-
(Gunsberger, 1958)	**Yugoslavia**	Case report	Case of pulmonary embolism	OSCM = 1	-	32	Iodized oil (not spec)	-	-	-

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Hohlbein, 1965)	**Germany**	Case report	Case of pulmonary embolism	OSCM = 1	Prim subfertility	26	40% Jodipin	20	Yes	-
(Levinson, 1963)	**USA**	Case report	Case of pulmonary embolism	OSCM = 1	Sec subfertility	49	Iodochlorol 27% iodine	-	-	-
(Elliott* et al., *1965)	**USA**	Case series	Cases of oil granuloma	OSCM = 3	-	-	OSCM	-	-	-
(Claus and Dochez, 1966)	**Belgium**	Case series	Cases of pulmonary embolism	OSCM = 2	-	27, 38	Lipiodol (not spec)	-	-	-
(Pear and Boyden, 1967)	**USA**	Case report	Case of retention of OSCM	OSCM = 1	-	28	Lipiodol (not spec)	-	-	-
(Aznar* et al., *1969)	**Mexico**	Case series	Cases of intravasation	OSCM = 141	Prim subfertility 21% Sec subfertility 18% Other indication 62%	26 – 30	Ioidized oil	2	Yes	Both tubes patent 40% Unilateral tubal occlusion 15% Bilateral tubal occlusion 30%
(Lau, 1969)	**Germany**	Case series	Cases of intravasation	OSCM = 1 WSCM = 2	-	24, 33, 35	Jodipin	-	-	Uncertain in all 3
(Bohm and Seewald, 1972)	**Germany**	Case series	Case series of PID/salpingitis	OSCM = 3	-	-	Jodipin	-	-	-

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Malter and Fox, 1972)	**USA**	Case report	Case of retention of OSCM	OSCM = 1	Prim subfertility	26	OSCM	-	-	-
(Charawanamuttu* et al., *1973)	**UK**	Case report	Case of retinal embolism	OSCM = 1	Sec subfertility	31	LipiodolÒ Ultra Fluid 40%	>20	Yes	Both tubes patent Uterine septum
(Weise* et al., *1973)	**Germany**	Case report	Case of oil granuloma	OSCM = 1	-	40	Jodipin	-	-	-
(Binder* et al., *1976)	**Romania**	Case report	Case of pulmonary embolism	OSCM = 1	Prim subfertility	26	Lipiodol (not spec)	-	No	Bilateral obstruction, unilateral hydrosalpinx, uterus arcuate.
(Bersi, 1977)	**Italy**	Case report	Case of oil granuloma	OSCM = 1	Prim subfertility, myomas	35	Lipiodol (not spec)	-	-	-
(Ishizuki* et al., *1992)	**Japan**	Case report	Case of thyroiditis	OSCM = 1	Subfertility (not spec). Graves' disease	28	Lipiodol (not spec)	-	-	-
(Dan* et al., *1990)	**Israel**	Case report	Case of pulmonary embolism with comatose state	OSCM = 1	Secondary subfertility	33	LipiodolÒ Ultra Fluid	6	-	Both tubes patent Bicornuate uterus
(Grosskinsky* et al., *1994)	**USA**	Case report	Case of oil granuloma	OSCM = 1	Secondary subfertility	29	Ethiodol	6	-	Both tubes patent

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Uzun* et al., *2004)	**Turkey**	Case report	Case of pulmonary embolism with comatose state	OSCM = 1	Secondary amenorrhoea	29	LipiodolÒ Ultra Fluid	-	-	Irregular endometrial cavity At least one tube patent
(Schutte* et al., *2006)	**Netherlands**	Case report	Case of retention of OSCM	OSCM = 1	Subfertility (not spec)	31	-	-	-	-
(Gotoh* et al., *2010)	**Japan**	Case report	Case of retention of OSCM	OSCM = 1	Subfertility (not spec)	39	OSCM	-	-	-
(Morii* et al., *2013)	**Japan**	Case report	Case of Fitz-Hugh-Curtis syndrome-like findings	OSCM = 1	Subfertility (not spec)	37	OSCM	-	-	-
(Omoto* et al., *2013)	**Japan**	Case report	Case of fetal goiter	OSCM = 1	Secondary subfertility	40	Lipiodol (not spec)	10	-	Both tubes patent
(Takeyama* et al., *2014)	**Japan**	Case report	Case of retention of OSCM	OSCM = 1	Subfertility (not spec)	34	OSCM	-	-	-
(Ueda* et al., *2016)	**Japan**	Case report	Case of pulmonary embolism	OSCM = 1	Subfertility (not spec), previous left tubectomy	27	LipiodolÒ Ultra Fluid	8	Yes	Left salpingectomy Right not described

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Ma* et al., *2016)	**China**	Case report	Case of maternal thyroid dysfunction	OSCM = 1	Subfertility (not spec)	33	OSCM	100	-	Both tubes patent
(Sasaki* et al., *2017)	**Japan**	Case report	Case of fetal goiter	OSCM = 1	Subfertility (not spec)	27	OSCM	-	-	-
(Li *et al.*, 2018)	**China**	Case series	Cases of (subclinical) hypothyroidism	OSCM = 14	-	-	OSCM	-	-	-
(Yamazaki *et al.*, 2019)	**Japan**	Case report	Case of fetal goiter	OSCM = 1	-	35	OSCM	3 HSGs	-	-
(Miyazaki et al., 2020)	**Japan**	Case report	Case of oil granuloma	OSCM = 1	Subfertility (not spec)	30	Lipiodol (not spec)	-	-	-
(Rubin, 1928)	**USA**	Cohort not specified	CO2-insufflation followed by HSG	OSCM = 66	-	-	Lipiodol (not spec)	-	-	Bilateral tubal occlusion 100%
(Witwer* et al., *1930)	**USA**	Retrospective cohort	Evaluate complications	OSCM = 512	-	-	Lipiodol (not spec)	4 (2.5 – 5.0)	Not standard	Not described. Only 1 case with tubal rupture in which bilateral hydrosalpinx
(Riche and Fayot, 1931)	**France**	Retrospective cohort	Evaluate complications	OSCM = 120	HSGs not for infertility 21%	-	Lipiodol (not spec)	-	-	-

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Schultze, 1932)	**Germany**	Retrospective cohort	Evaluate complications	OSCM = 600	-	-	Jodipin 40%	-	-	-
(Volk, 1936)	**Germany**	Retrospective cohort	Evaluate complications	OSCM = 316	Not reported. Also for other indications	-	OSCM	-	-	-
(Nordio, 1938)	**Italy**	Retrospective cohort	Evaluate intravasation	OSCM = 106	-	-	Iodized oil	-	-	-
(Feiner, 1942)	**USA**	Cohort not specified	Cases of PID/ salpingitis/ death	OSCM = 337	-	-	Lipiodol (not spec)	-	-	Bilateral tubal occlusion 44% Uni/bilateral patent tubes 56%
(Madsen, 1942)	**Denmark**	Retrospective cohort	Evaluate complications	OSCM = 490	Complicated cases, 2/6 for subfertility	Complicated cases: 45 (30-56)	Complicated cases: 5/6 Jodipin 20%, 1/6 Jodumbrin	Complicated cases: 5.3 (3.5-9)	No	Complicated cases: 4/6 bilateral tubal occlusion
(Rutherford, 1948)	**USA**	Cohort not spec	Description of technique, HSG after CO2-insufflation	OSCM = 417	Prim subfertility 75% Sec subfertility 25%	19 – 41	Lipiodal	Max 12	Yes	After failed CO2-insufflation, all bilateral tubal occlusion

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Bang, 1950)	**Denmark**	Retrospective cohort	Evaluate complications	OSCM = 900	Prim/sec subfertility 98% Referred for sterilisation 1.1%	-	Half Jodubrin, half Jodipin, 15 cases Lipiodol	-	-	-
(Vara, 1950)	**Finland**	Retrospective cohort	Evaluate intravasation	OSCM = 1119	Prim/sec subfertility	-	Jodipin, Lipiodol, Jodolja, Neo-Hydrol (20-40%)	-	No	Bilateral patent tubes 53%
(Bergin, 1951)	**UK**	Retrospective cohort	Evaluate complications	OSCM = 201 WSCM = 69	-	-	Lipiodol (not spec)	-	-	-
(Drukman and Rozin, 1951)	**Israel**	Retrospective cohort	Evaluate intravasation	OSCM = 2000	-	-	Lipiodol 40% 'Assia' and 'Laffay'	-	-	-
(Robins and Shapira, 1951)	**USA**	Retrospective cohort	Evaluate HGS and complications	OSCM = 4800	Not reported. Also for other indications	-	Lipiodol Lafay 28%	-	-	-

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Kika, 1954)	**Japan**	Retrospective cohort	Evaluate intravasation and tuberculosis	OSCM = 1200	Complicated cases: Prim subfertility 90% Sec subfertility 10%	-	20-40% Moljodol	4-8	Yes	Complicated cases: Bilateral patent tubes 11% Bilateral or unilateral tubal occlusion 89%
(Zachariae, 1955)	**Denmark**	Cohort not spec	Evaluate complications	OSCM = 505 WSCM = 55	Complicated cases: Prim subfertility 67% Non-infertility 33%	Complicated cases: 25 – 59	Various OSCM (esp iodumbrin, neohydriol)	-	Yes	-
(Norris, 1956)	**Canada**	Cohort not spec	Evaluate complications	OSCM = 961	Subfertility (not spec)	-	Lipiodol (not spec)	3.5-4 (max 10)	-	Bilateral tubal occlusion 35%
(Stoll and Zeitz, 1956)	**Germany**	Retrospective cohort	Evaluate intravasation	OSCM = 2236	-	-	Jodipin	-	-	-
(Frischkorn, 1958)	**Germany**	Retrospective cohort	Evaluate intravasation	OSCM = 281 WSCM = 52	-	-	Jodipin 40%	8	-	-
(Woltz* et al., *1958)	**USA**	Cohort not spec	Evaluate complications	OSCM = 500	-	-	Iodochlorol	Average 7-10 (3-20)	-	Bilateral tubal occlusion 12% Hydrosalpinx 4.2%

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Palmer, 1960)	**USA**	Cohort not spec	Evaluate pregnancy rate and complications	OSCM = 258	Subfertility (not spec)	-	Ethiodol	-	-	-
(Kuzavova, 1964)	**Russia**	Retrospective cohort	Evaluate intravasation after HSG	OSCM = 730	Only intravasation cases: all subfertility	Intravasation cases: 20 – 37	Lipiodol (not spec)	3-5	No	Cases with intravasation: Bilateral tubal occlusion 54%
(Heinen and Schussler, 1966)	**Germany**	Retrospective cohort	Evaluate intravasation after HSG	OSCM = 122	-	-	Jodolen	-	-	-
(Geary* et al., *1969)	**USA**	Retrospective cohort	Evaluate pregnancy rates and complications	OSCM = 501	Prim subfertility 55% Sec subfert 29% Other subfertility: 16%	-	Lipiodol (not spec)	10	-	Both tubes patent 79% Unilateral tubal occlusion 7.4% Bilateral tubal occlusion 8.8%
(Mackey* et al., *1971)	**USA**	Retrospective cohort	Evaluate pregnancy rates	OSCM = 221 WSCM = 63	OSCM: Prim subfertility 63% Sec subfertility 37%	OSCM: 27.3 – 28.9	Ethiodol	-	Yes	-

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Buytaert and Meulyzer, 1977)	**Belgium**	Retrospective cohort	Evaluate pregnancy rates and complications	OSCM = 208	Prim subfertility 65% Sec subfertility 35%	-	Lipiodol 40%	3-6	Yes	-
(Bateman* et al., *1980)	**USA**	Cohort not spec	Evaluate complications after HSG	OSCM = 533	Subfertility and preoperative HSG before tuba reconstruction	Intravasation cases: 24-33	Ethiodol (37%)	1.5 – 12	Yes	Cases with complications: Bilateral tubal occlusion 69%
(La Sala* et al., *1982)	**Italy**	Retrospective cohort study	Evaluate intravasation after HSG	OSCM = 119	Sterility, suspected uterine malformations and other pathologies	-	LipiodolÒ ultra fluid (40%)	5-10	Yes	-
(Rasmussen* et al., *1987)	**Denmark**	Cohort not spec	Evaluate pregnancy rates after HSG	OSCM = 294	Prim subfertility 59% Sec subfertility 41%	28 (19 – 40)	LipiodolÒ Ultra Fluid	5-10	Yes	Both tubes patent: 41% Unilateral tubal occlusion 14% Bilateral tubal occlusion 19%

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Nunley* et al., *1987)	**USA**	Cohort not spec	Evaluate intravasation	OSCM = 593	Subfertility (not spec) Complications: 11 preoperative before tubal reconstruction	-	Ethiodol	2.0-10	Yes	Both tubes patent 41% Unilateral tubal occlusion 9.8% Bilateral tubal occlusion 44%
(Barqawi* et al., *2007)	**Jordan**	Retrospective cohort study	Compare pregnancy between OSCM and WSCM	OSCM = 35 WSCM = 40	Subfertility (not spec)	OSCM 28 (3)	OSCM	10-20	Yes	None bilateral tubal occlusion
(Mekaru* et al., *2008)	**Japan**	Cohort not spec	Evaluate thyroid function after HSG	OSCM = 220	Subfertility (not spec)	35 (4.6)	Lipiodol (not spec)	Average 5-10	Yes	-
(Liu* et al., *2010)	**China**	Retrospective cohort	Compare complications after HSG	OSCM = 100 WSCM = 300	-	Mean 31 (25-37)	Ioidized oil (not spec).	5-10	Yes	OSCM: Bilateral tubal occlusion 12% Unilateral tubal occlusion 15%

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Satoh* et al., *2015)	**Japan**	Retrospective cohort study	Evaluate neonatal thyroid function	OSCM = 212	-	-	Lipiodol (480)	5-10	-	-
(So* et al., *2017)	**Japan**	Cohort not spec	Compare maternal thyroid function after HSG with OSCM and WSCM	OSCM = 164 WSCM = 94	OSCM: Sec subfertility 36.6% < 3 months before HSG normal thyroid function	34 (4.3)	Lipiodol (480)	5	-	OSCM: Bilateral tubal occlusion 11%
(Brown* et al., *1949)	**USA**	Cohort not spec	Evaluate absorption of OSCM	OSCM = 118	-	-	Lipiodol (not spec) Iodochlorol Lipoiodine	8-12	No	-
(Slater* et al., *1959)	**USA**	Prospective study	Evaluate thyroid function after HSG with OSCM and WSCM	OSCM = 10 WSCM = 18	All clinically euthyroid.	-	Ethiodol	-	-	-
(Brent* et al., *2006)	**Australia**	Prospective cohort	Evaluate pregnancy rates and complications after HSG	OSCM = 100	Prim subfertility 60% Sec subfertility 40%	36.83 (4.26)	LipiodolÒ Ultra Fluid	10 (sometimes >10)	Yes	Bilateral tubal patency 80% Bilateral tubal occlusion 4%

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Kaneshige* et al., *2015)	**Japan**	Prospective cohort	Evaluate maternal thyroid function after HSG	OSCM = 22	Subfertility (not spec)	36+2.45	Lipiodol (not spec)	6.1 (4.0-9.0)	-	None bilateral tubal occlusion
(Tan et al., 2019)	**China**	Prospective cohort	Evaluate image quality and complications	OSCM = 165 WSCM = 63	OSCM: Prim subfertility 64%	OSCM 31.36 (4.99)	Ethiodized poppyseed oil	OSCM 6-8	Yes	-
(Schwabe* et al., *1983)	**USA**	RCT	Compare pregnancy rates after HSG	OSCM = 56 WSCM = 65	Subfertility (not spec)	-	Ethiodol	-	Yes	-
(Alper* et al., *1986)	**Canada**	RCT	Compare pregnancy rates after HSG	OSCM = 58 WSCM = 73	Subfertility (not spec)	OSCM 29 (2.9)	LipiodolÒ Ultra Fluid	10.5 (SD 4.7)	Yes	None bilateral tubal occlusion
(Lindequist* et al., *1991)	**Denmark**	RCT	Evaluate complications after HSG	OSCM = 103 WSCM = 314	Subfertility (not spec)	-	LipiodolÒ Ultra Fluid	5-10	Yes	OSCM: Bilateral tubal patency 47%
(Lindequist* et al., *1994)	**Denmark**	RCT	Compare pregnancy rates and complications after HSG	OSCM = 123 WSCM = 122	OSCM: Prim subfertility 60% Sec subfertility 40%	OSCM 29.9 (21-43)	LipiodolÒ Ultra Fluid	5-10	Yes	OSCM: Bilateral tubal patency 54%
(Nugent* et al., *2002)	**UK**	RCT	Evaluate pregnancy rates after HSG	OSCM = 17	Prim subfertility 71% Sec subfertility 29%	31 (1.1)	Lipiodol (not spec)	5.8 (0.7)	Yes	All patent

Supplementary Table 5. (Continued)

							Intervention			
Reference	Country	Study design	Purpose	**No.**	**Population**	**Mean age at HSG (y) (SD or IQR) Median (range)**	**Contrast**	**Volume OSCM (mL) Median (range)**	**Fluoroscopy**	**HSG results**
(Steiner* et al., *2003)	**USA**	RCT	Evaluate pregnancy rates after HSG	OSCM = 30	All subfertility Prim subfertility 47%	33 (3.6)	Ethiodol	5-10	Yes	All patent
(Johnson* et al., *2004)	**Australia**	RCT	Evaluate pregnancy rates after	OSCM = 73	All subfertility Prim subfertility 55%	33.9 (2.9)	LipiodolÒ Ultra Fluid	10	Yes	Bilateral tubal occlusion 8.3%
(Dreyer* et al., *2017)	**Netherlands**	RCT	Compare pregnancy rates after HSG	OSCM = 550 WSCM = 556	All subfertility Prim subfertility OSCM 67.3%	OSCM 32.8 (30.1-35.7)	LipiodolÒ Ultra Fluid.	9.0 (5.7-15.0)	Yes	OSCM: Bilateral tubal patency 86% Bilateral tubal occlusion 1.6%

Not spec = not specified. OSCM = oil-soluble contrast media. WSCM = water-soluble contrast media.

Supplementary Table 6. Risk of bias assessment

First author, year of publication	Study type	1. Was the study's target population a close representation of the national population in relation to relevant variables	2. Was the sampling frame a true or close representation of the target population?	3. Was some form of random selection used to select the sample, OR, was a census undertaken?	4. Was the likelihood of non-response bias minimal?	5. Were data collected directly from the subjects (as opposed to a proxy)?	6. Was an acceptable case definition used in the study?	7. Was the study instrument that measured the parameter of interest shown to have reliability and validity (if necessary)?	8. Was the same mode of data collection used for all subjects?	9. Were the numerator(s) and denominator(s) for the parameter of interest appropriate?	10. Summary of the overall risk of study bias:	0-3 Low risk, 4-6 moderate risk, 7-9 High risk.
Alper, 1986	RCT	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	High risk	Low risk	Low risk	1	Low risk
Bang, 1950	Retrospective cohort	Low risk	Low risk	Unclear	Low risk	Unclear	High risk	High risk	Unclear	Low risk	5	Moderate risk
Barqawi, 2007	Retrospective cohort	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	High risk	Low risk	Low risk	1	Low risk
Bateman, 1980	Cohort not spec	High risk	Low risk	Low risk	Low risk	Low risk	Unclear	High risk	Unclear	Low risk	4	Moderate risk
Bergin, 1951	Retrospective cohort	Unclear	Unclear	Unclear	Unclear	Unclear	High risk	High risk	Unclear	Low risk	8	High risk
Brent, 2006	Prospective cohort	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	High risk	Low risk	Low risk	1	Low risk
Brown, 1949	Cohort not spec	Unclear	Unclear	Unclear	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	6	Moderate risk
Buytaert, 1977	Retrospective cohort	Low risk	Unclear	Unclear	Low risk	Unclear	High risk	High risk	Unclear	Low risk	6	Moderate risk
Dreyer, 2017	RCT	Low risk	Low risk	Low risk	Low risk	Low risk	High risk	High risk	Low risk	Low risk	2	Low risk
Drukman, 1951	Retrospective cohort	High risk	Unclear	Unclear	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	6	Moderate risk

Supplementary Table 6. (Continued)

First author, year of publication	Study type	1. Was the study's target population a close representation of the national population in relation to relevant variables	2. Was the sampling frame a true or close representation of the target population?	3. Was some form of random selection used to select the sample, OR, was a census undertaken?	4. Was the likelihood of non-response bias minimal?	5. Were data collected directly from the subjects (as opposed to a proxy)?	6. Was an acceptable case definition used in the study?	7. Was the study instrument that measured the parameter of interest shown to have reliability and validity (if necessary)?	8. Was the same mode of data collection used for all subjects?	9. Were the numerator(s) and denominator(s) for the parameter of interest appropriate?	10. Summary of the overall risk of study bias:	0-3 Low risk, 4-6 moderate risk, 7-9 High risk.
Feiner, 1942	Retrospective cohort	Low risk	Unclear	Unclear	Low risk	High risk	High risk	High risk	Unclear	High risk	6	Moderate risk
Frischkorn, 1958	Retrospective cohort	Unclear	Unclear	Low risk	Low risk	Unclear	High risk	High risk	Unclear	Low risk	6	Moderate risk
Geary, 1969	Retrospective cohort	High risk	Low risk	Low risk	Low risk	Unclear	High risk	High risk	Unclear	Low risk	5	Moderate risk
Heinen, 1966	Retrospective cohort	Unclear	Unclear	Unclear	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	6	Moderate risk
Johnson, 2004	RCT	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	High risk	Low risk	Low risk	1	Low risk
Kaneshige, 2015	Prospective cohort	Low risk	Unclear	Unclear	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	2	Low risk
Kika, 1954	Retrospective cohort	Low risk	High risk	Unclear	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	4	Moderate risk
Kuzavova, 1964	retrospective cohort	Low risk	Low risk	Low risk	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	3	Low risk
La Sala, 1982	Retrospective cohort	High risk	Low risk	Low risk	Low risk	Unclear	High risk	High risk	Unclear	Low risk	5	Moderate risk
Lindequist, 1991	RCT	High risk	Low risk	Low risk	Unclear	Low risk	Low risk	High risk	Low risk	Low risk	3	Low risk

Supplementary Table 6. (Continued)

First author, year of publication	Study type	1. Was the study's target population a close representation of the national population in relation to relevant variables	2. Was the sampling frame a true or close representation of the target population?	3. Was some form of random selection used to select the sample, OR, was a census undertaken?	4. Was the likelihood of non-response bias minimal?	5. Were data collected directly from the subjects (as opposed to a proxy)?	6. Was an acceptable case definition used in the study?	7. Was the study instrument that measured the parameter of interest shown to have reliability and validity (if necessary)?	8. Was the same mode of data collection used for all subjects?	9. Were the numerator(s) and denominator(s) for the parameter of interest appropriate?	10. Summary of the overall risk of study bias:	0-3 Low risk, 4-6 moderate risk, 7-9 High risk.
Lindequist, 1994	RCT	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	High risk	Low risk	Low risk	1	Low risk
Liu, 2010	Retrospective cohort	Low risk	Low risk	Low risk	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	3	Low risk
Mackey, 1971	Retrospective cohort	Low risk	Low risk	Low risk	Low risk	High risk	High risk	High risk	High risk	Low risk	4	Moderate risk
Madsen, 1942	Retrospective cohort	High risk	Low risk	Low risk	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	4	Moderate risk
Mekaru, 2008	Cohort not spec	Low risk	Low risk	Unclear	Low risk	Low risk	Low risk	Low risk	High risk	Low risk	2	Low risk
Nordio, 1938	Retrospective, cohort	Unclear	Unclear	Unclear	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	6	Moderate risk
Norris, 1956	Cohort not spec	Low risk	Low risk	Low risk	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	3	Low risk
Nugent, 2002	RCT	Low risk	Low risk	Low risk	Low risk	Unclear	High risk	High risk	Unclear	Low risk	4	Moderate risk
Nunley, 1987	Cohort not spec	Low risk	High risk	Low risk	Low risk	Unclear	Unclear	High risk	Unclear	Low risk	5	Moderate risk
Palmer, 1960	Cohort not spec	Low risk	Unclear	Unclear	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	5	Moderate risk

Supplementary Table 6. (Continued)

First author, year of publication	Study type	1. Was the study's target population a close representation of the national population in relation to relevant variables	2. Was the sampling frame a true or close representation of the target population?	3. Was some form of random selection used to select the sample, OR, was a census undertaken?	4. Was the likelihood of non-response bias minimal?	5. Were data collected directly from the subjects (as opposed to a proxy)?	6. Was an acceptable case definition used in the study?	7. Was the study instrument that measured the parameter of interest shown to have reliability and validity (if necessary)?	8. Was the same mode of data collection used for all subjects?	9. Were the numerator(s) and denominator(s) for the parameter of interest appropriate?	10. Summary of the overall risk of study bias:	0-3 Low risk, 4-6 moderate risk, 7-9 High risk.
Rasmussen, 1987	Cohort not spec	Low risk	Low risk	High risk	Low risk	Unclear	High risk	High risk	Unclear	Low risk	5	Moderate risk
Riche, 1931	Retrospective cohort	High risk	Low risk	Low risk	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	4	Moderate risk
Robins, 1951	Cohort not spec	Unclear	Unclear	Low risk	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	5	Moderate risk
Rubin, 1928	Cohort not spec	Unclear	High risk	Unclear	Low risk	Unclear	High risk	High risk	Unclear	Unclear	8	High risk
Rutherford, 1948	Cohort not spec	Low risk	High risk	Low risk	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	4	Moderate risk
Satoh, 2015	Retrospective cohort	Unclear	Unclear	Low risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk	3	Low risk
Schultze, 1932	Retrospective cohort	Unclear	Low risk	Low risk	Low risk	Unclear	High risk	High risk	Unclear	Low risk	5	Moderate risk
Schwabe, 1983	RCT	Low risk	Unclear	Low risk	Low risk	Unclear	High risk	High risk	Unclear	Low risk	5	Moderate risk
Shuhei So, 2017	Cohort not spec	Low risk	Unclear	Unclear	High risk	Low risk	Low risk	Low risk	Low risk	Low risk	3	Low risk
Slater, 1959	Prospective cohort	High risk	Unclear	Unclear	Low risk	Low risk	Unclear	Unclear	Low risk	High risk	6	Moderate risk

Supplementary Table 6. (Continued)

First author, year of publication	Study type	1. Was the study's target population a close representation of the national population in relation to relevant variables	2. Was the sampling frame a true or close representation of the target population?	3. Was some form of random selection used to select the sample, OR, was a census undertaken?	4. Was the likelihood of non-response bias minimal?	5. Were data collected directly from the subjects (as opposed to a proxy)?	6. Was an acceptable case definition used in the study?	7. Was the study instrument that measured the parameter of interest shown to have reliability and validity (if necessary)?	8. Was the same mode of data collection used for all subjects?	9. Were the numerator(s) and denominator(s) for the parameter of interest appropriate?	10. Summary of the overall risk of study bias:	0-3 Low risk, 4-6 moderate risk, 7-9 High risk.
Steiner, 2003	RCT	Low risk	Low risk	Low risk	Low risk	Unclear	High risk	High risk	Unclear	Low risk	4	Moderate risk
Stoll, 1956	Retrospective cohort	Unclear	Unclear	Low risk	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	5	Moderate risk
Tan, 2019	Prospective cohort	Low risk	Low risk	Moderate risk	Low risk	Low risk	Low risk	High risk	Low risk	Low risk	2	Low risk
Vara, 1950	Retrospective cohort	Low risk	Unclear	Unclear	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	5	Moderate risk
Volk, 1936	Retrospective cohort	High risk	Low risk	Low risk	Low risk	Unclear	High risk	High risk	Unclear	Unclear	6	Moderate risk
Witwer, 1930	Retrospective cohort	Unclear	Unclear	Unclear	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	6	Moderate risk
Woltz, 1958	Cohort not spec	Unclear	High risk	Low risk	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	5	Moderate risk
Zachariae, 1955	Cohort not spec	High risk	Low risk	Low risk	Low risk	Unclear	Low risk	High risk	Unclear	Low risk	4	Moderate risk

CHAPTER 6

Virtual Reality as pain relief during hysterosalpingography: a randomized controlled trial

K. Rosielle, A. P. van Haaps, E.A.M. Kuijper, N. Tonch, D.E.N.B. Karim, M.M. Oskam, R. van den IJssel, B.W.J. Mol, C.B. Lambalk, K. Dreyer, V. Mijatovic

Manuscript in preparation

ABSTRACT

Study question
Is Virtual Reality an effective non-pharmacological tool to reduce procedural pain during hysterosalpingography (HSG)?

Summary answer
An HSG with Virtual Reality does not result in lower pain scores than an HSG without Virtual Reality.

What is known already
An HSG is often rated as uncomfortable or painful, with a median pain score of 5, measured using the Visual Analogue Scale (VAS, scale 0.0-10.0cm). Virtual Reality has been proven successful to reduce acute procedural pain during different medical procedures and interventions.

Study design, size, duration
We performed a two-centre open label randomized controlled trial between January 2021 and October 2022. After informed consent, women were randomized between HSG with or without Virtual Reality headset during HSG. Due to the nature of the intervention, the study was not blinded. The study was registered in the Netherlands Trial register with registration number NL9203.

Participants/materials, setting, methods
The primary endpoint was procedural pain, measured using VAS (scale 0.0 – 10.0cm). Procedural pain consisted of overall pain score and peak pain score during the procedure measured immediately after the HSG procedure. Secondary endpoints included use of analgesics after the procedure, satisfaction, Virtual Reality preferences and side effects.

Main results and the role of chance
We randomized a total of 134 women, 69 to the intervention group and 65 to the control group. The mean VAS for peak pain was 6.8 in the intervention group versus 6.6 in the control group (p-value 0.574). The mean VAS for overall pain was 5.0 in the intervention group versus 4.9 in the control group (p-value 0.915). The use of analgesics within the first 24 hours after the HSG was comparable between the two groups (18.8% in the intervention group versus 21.3% in the control group). Satisfaction scores were also comparable. There were no differences in the prevalence of symptoms that are associated as side effects of the use of Virtual Reality. The mean score for being distracted by Virtual Reality was 3.54 on a scale of 1-5.

Limitations, reasons for caution

The study was not blinded. Reasons for study decline were anxiety or wanting full control during HSG, which might have created selection bias. The distraction score possibly indicates that the level of immersiveness during Virtual Reality was not optimal due to the lack of sound and/or the type of VR applications. Future studies should investigate whether more immersive or interactive Virtual Reality applications decrease pain scores during HSG.

Wider implications of the findings

Virtual Reality does not reduce pain during HSG, and should therefore not be used.

Study funding/competing interest(s)

All authors have no conflict of interest related to this manuscript. The department of reproductive medicine of the Amsterdam UMC location VUmc has received several research and educational grants from Guerbet, Merck and Ferring not related to the submitted work.

Trial registration number

The trial was registered prospectively in the Netherlands Trial Register (trialregister.nl, currently accessible on trialsearch.who.int), with corresponding registration number NL9203.

Trial registration date

16-01-2021

Date of first patient's enrolment

The first participant was enrolled on January 19th 2021.

INTRODUCTION

Infertility affects 11.4% of women wishing to conceive in the Netherlands and can compromise quality of life significantly (1, 2). Among the most common causes of infertility are tubal pathology and pelvic adhesions; therefore evaluation of the Fallopian tubes is part of the fertility work-up (3, 4). The traditional method for evaluation of the Fallopian tubes during fertility work-up is hysterosalpingography (HSG). During HSG, iodine-containing contrast is infused into the uterine cavity and subsequently into the Fallopian tubes. The contrast is visualized using fluoroscopy guiding to determine whether the Fallopian tubes are patent. The HSG was initially introduced as a diagnostic test. However, it has also been proven to have a therapeutic effect in women with unexplained infertility, especially when oil-based contrast is used (5-7).

HSG as a tubal patency test can cause discomfort and pain during and immediately after the procedure, with average reported pain scores ranging from 3.7 to 5.0 on the Visual Analogue Scale (VAS, scale 0.0-10.0cm) (8-10). The painful cramping sensations are caused by cannulation of the cervix with instruments, distention of the uterine cavity due to infusion of contrast, increase of pressure in the Fallopian tubes and peritoneal irritation caused by contrast spillage from the Fallopian tubes (11). The different parts of the female reproductive system are innervated by different pathways and nerves, making it difficult to target the overall pain with local analgesics (12).

Several regimens of analgesics have been studied using various routes of administration, i.e. oral, topical (a spray on the cervix or intrauterine installation with analgesics), locally injected and intravenous administration of analgesics. A meta-analysis by Hindocha et al. showed that topical and locally injected analgesics can reduce the median pain score during the procedure (mean difference (MD) -0.63, 95%CI -1.06 to -0.19 after topical analgesics and MD -1.31, 95% CI -1.55 to -1.07 after locally injected analgesics), although the quality of evidence was low to very low (11). The most obvious pain reduction was established by intravenous opioids (MD −3.53, 95% CI −4.29 to −2.77). However, because of potential cardiorespiratory side effects, the administration of intravenous opioids requires continuous monitoring, making this a less attractive mode of pain management to apply during HSG.

Virtual Reality (VR) has recently been introduced in medical practice and can be used to reduce pain and anxiety in a non-pharmacological way. It creates a distraction from the present and diverts attention away from physical sensations by taking the user into a three-dimensional virtual world, usually without any major side effects (13). Virtual Reality for medical use has been proven successful in decreasing acute pain during different medical procedures and interventions, including a reduction in procedural pain

during hysteroscopy and episiotomy repair (14-16). Aside from targeting pain, VR can also lead to lower stress scores (17) and anxiety scores (Hoxhallari et al., 2019) during medical procedures. When used in a perioperative setting, VR is associated with a lower stress score (18) and higher satisfaction score (Haisley et al., 2020).

Considering these favorable outcomes of VR on procedural pain, it is expected that the application of VR could lead to a reduction in pain during HSG. Therefore, the aim of this study is to investigate whether VR is an effective non-pharmacological method to reduce pain during HSG in infertile women undergoing infertility workup.

MATERIALS AND METHODS

Study design

This randomized controlled trial was performed in one university hospital, with two locations, and one teaching hospital in the Netherlands. The study was approved by the institutional review board of the Amsterdam UMC, location VUmc (reference number 2020.0687). The trial was prospectively registered in the Dutch trial register (www.trialregister.nl; registration number NL9203).

Participants

All women referred for HSG were screened for eligibility. Women were eligible for study participation if they were scheduled for an HSG during infertility workup and had sufficient knowledge of the Dutch or English language to read and understand the patient information sheet. Women were excluded from study participation if they had a history of surgery to the cervix, if they had had an HSG before, if they had a known allergy to iodinated contrast and if they currently used antidepressants, sedatives or analgesics on a daily basis.

Randomization

After providing written informed consent, participants were randomized between HSG with VR (intervention group) or HSG without VR (control group). Randomization was performed in a 1:1 ratio with variable block sizes varying from 4-8 using the web-based system Castor (Castor Electronic Data Capture, Ciwit BV, Amsterdam, the Netherlands) and stratified per centre. Due to the nature of the intervention, blinding of participants and physicians was not possible.

Study procedures

The HSG was performed according to local standard procedure protocols. Prior to the procedure, all women received oral and written information on the HSG procedure.

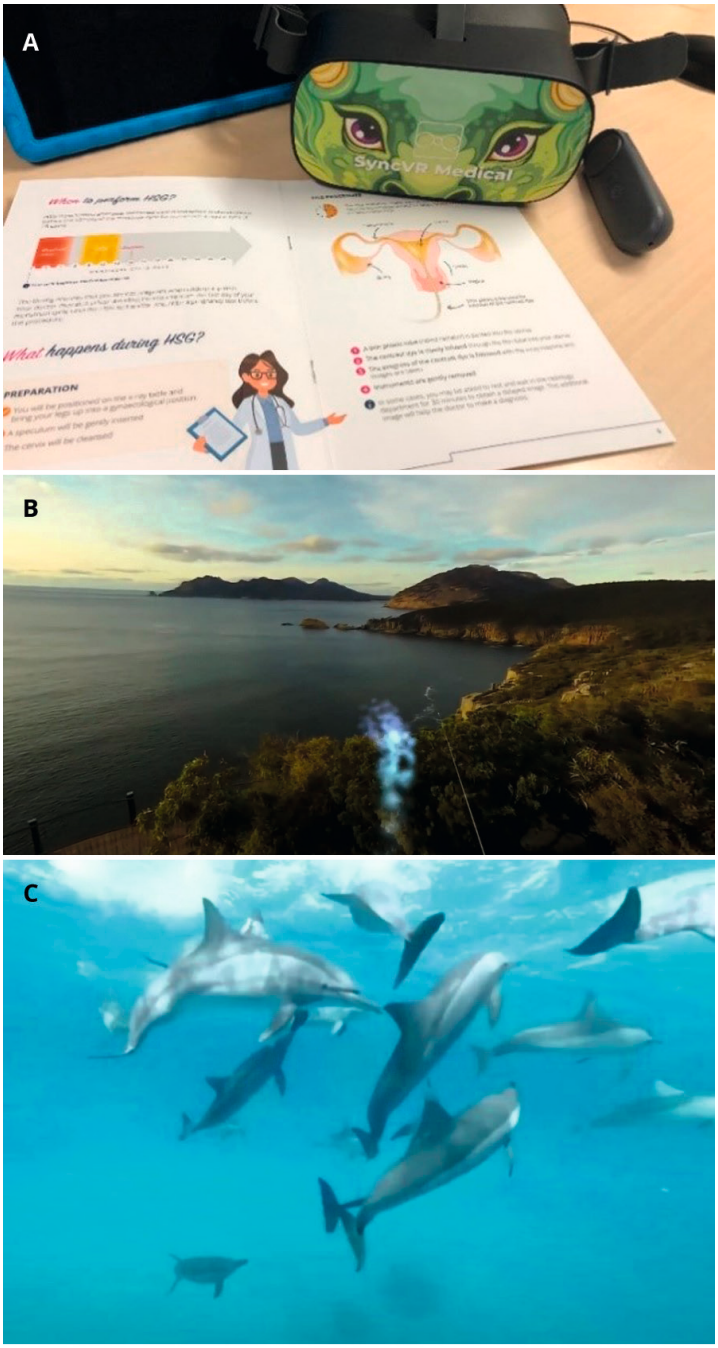
An HSG was scheduled in the follicular phase of the menstrual cycle after cessation of menstrual bleeding, or after cessation of progesterone induced bleeding in case of amenorrhea. Women were advised to take non-steroid anti-inflammatory drugs such as naproxen or ibuprofen as pain management the evening before and 2 hours before the HSG. During the procedure, the patient was placed in lithotomy position after which a speculum was inserted. The cervix was cannulated with a cervical vacuum cup, a metal cannula (hysterophore) or an HSG balloon catheter, depending on the preference of the physician performing the HSG and local availability. During the infusion of approximately 10 ml of oil-based (Lipiodol Ultra Fluid®, Guerbet, Villepinte, France) or water-based contrast medium (iodixanol, Visipaque®, General Electric Healthcare, Buc, France), four to six radiographs were obtained. The radiographs were examined by a gynaecologist, fertility doctor and/or radiologist to evaluate the patency of both Fallopian tubes.

Participants assigned to the intervention group were given a short instruction of the VR headset prior to the HSG procedure. After meeting their physician and getting into lithotomy position on the radiology table, women placed the VR headset over their eyes. Participants were asked to wear the VR headset during the entire procedure, until all instruments were removed. Sound from the VR headset was switched off to enable participants to communicate with the physician performing the procedure at all times. A dedicated researcher was present during the entire procedure for technical support with the VR system. In case of complaints or potential side-effects of the VR headset, the participants were able to adjust or remove the headset themselves or ask for assistance from the dedicated researcher.

Virtual Reality Technology

The VR headset and software used in this study were CE-certified and obtained from SyncVR (SyncVR Medical B.V., Utrecht, the Netherlands). The head-mounted PICO G2 4K device (Pico Interactive Inc., San Francisco, United States of America) included a VR headset and a remote control (Figure 1a). Participants were able to navigate the VR headset on their own using the remote control. However, if they felt unsure navigating the system the researcher was able to take over control of the VR headset on a hand-held device. SyncVR Relax & Distract is a software module installed on the VR headset which enables patients to choose between approximately twenty relaxing movies and breathing exercises, all designed for use in medical practice (Figure 1b,c). Games, requiring movement of the participant, and movies with an element of surprise were removed from the library to decrease the chance of sudden movements during the HSG procedure. It was thought that sudden movements were undesirable and unsafe for participants (Figure 1a-c).

Figure 1a. The head-mounted VR device (PICO G2 4K) with remote control, **b.** example of the nature movies, **c.** example of the breathing exercises



Outcomes

The primary outcome of this study was pain experience during the HSG procedure, defined as the overall pain and the most severe pain experienced. Pain was measured using a VAS ruler (scale 0.0 - 10.0cm). Secondary outcomes included satisfaction, usage of oral analgesics following HSG, preference of type of VR distraction, and side effects. The anxiety level and the VR expectations prior to HSG were reported through a questionnaire, since these were considered possible confounding factors.

Data collection

First, all participants received a short questionnaire with an adapted version of the Amsterdam Preoperative Anxiety and Information Scale (APAIS; N. Moerman, F.S.A.M. van Dam, M.J. Muller and J. Oosting) to determine their level of anxiety while waiting for the procedure. It additionally contained questions on their expectations and expected preferences of the VR headset and questions on their usage of oral analgesics prior to the procedure were asked. The questions on VR software preference included open field questions to explore the rationale for preferences.

Immediately after the procedure, participants were asked to score their highest pain level and their overall pain level during HSG using a VAS-ruler (0.0 - 10.0cm). The second questionnaire was completed within 30 minutes after the HSG and contained questions on possible side effects of the use of VR and participant satisfaction with the procedure (5-point Likert scale). In addition, participants in the intervention group received questions on their satisfaction with their chosen VR application.

Finally, the third and last questionnaire was sent to all participants by email using Castor EDC, a day after the procedure. This last questionnaire inquired on side effects or complaints, willingness to undergo the procedure again and use of oral analgesics in the first 24 hours after the procedure. All questionnaires can be found in the appendix S3. A subgroup of patients received a more elaborate questionnaire before and immediately after the procedure with questions on their expectations of VR and their experiences.

Medical details from the HSG procedure (for example the type and volume of contrast medium used, instrument used for instalment of contrast) and results of the HSG were collected from the patient files.

Statistics

All data were analysed according to the intention-to-treat principle using IBM SPSS for Windows, version 26.0 (IBM Corp., Armonk, NY, USA). Baseline characteristics were presented using descriptive statistics with median and interquartile range for non-

normally distributed continuous variables, and with mean and SD for normally distributed continuous variables. Categorical variables were presented as absolute numbers and percentages. Continuous outcomes were analysed with the use of an independent t-test or the Mann-Whitney U-test as appropriate, whereas categorical variables were analysed with use of the Chi-square test or the Fishers exact test. Two-sided P values of less than 0.05 were considered to indicate statistical significance. A multiple linear regression model was fitted to assess the influence of predefined confounders on the primary outcome measures. Anxiety level, VR expectations and use of analgesics were analysed as possible confounders prior to the procedure. The used instruments during HSG, type of contrast used and HSG outcomes were analysed as possible confounders during the HSG procedure.

Sample size

Based on previous research in a similar, Dutch population, we anticipated an overall pain score of 5.0 cm (in VAS, scale 0.0 - 10.0 cm) in the control group (5). To detect a difference of 1.5cm in pain score, we calculated that the trial needed a sample size of 134 women (67 in each group) to obtain a power of 80%, with an alpha of 5% and a loss to follow-up rate of 5% using PASS (15.0.5).

RESULTS

Between January 2021 and October 2022, 211 women were screened for eligibility, of which 185 were eligible and 135 provided informed consent. We randomly allocated 70 women to HSG with VR (intervention group), and 65 to HSG without VR (control group). One woman from the intervention group did not undergo HSG after randomization and was considered a drop-out (Figure 2).

Randomization led to two equal groups in terms of baseline characteristics. The baseline characteristics are shown in Table 1.

Figure 2. Flowchart of inclusions

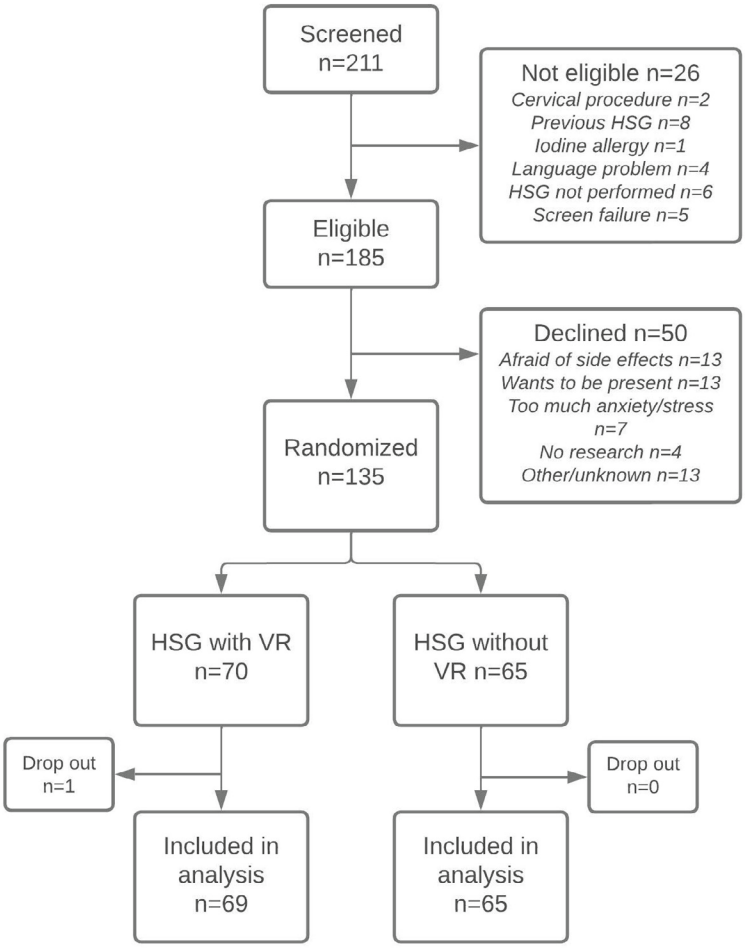


Table 1. Baseline characteristics.

Characteristic	Intervention n= 69	Control n= 65
Age (year, median (IQR))	35 (5.00)	35 (4.00)
BMI (kg/m^2, mean (st dev))	23.9 (4.4)	24.6 (4.4)
Duration of infertility (months, mean (st dev))	29.8 (15.4)	27.0 (16.4)
Primary infertility (%)	47 (72.3)	45 (65.2)
Risk assessment for tubal pathology* (%)		
High risk	11 (15.9)	12 (18.5)
Low Risk	58 (84.1)	53 (81.5)
Cause of infertility (%)		
Male factor	7 (10.1)	5 (7.7)
Female factor	12 (17.4)	7 (10.8)
Male and female factor	1 (1.4)	2 (3.1)
Unexplained	49 71.0)	51 (78.5)
Current smoker# (%)	3 (4.3)	4 (6.3)

*Considered as high risk for tubal pathology were history of symptomatic Chlamydia Trachomatis infection, a positive Chlamydia Antibody Test (CAT) on screening, history of Pelvic Inflammatory Disease (PID), history of endometriosis, history of surgery on ovaries or fallopian tubes and a history of ectopic pregnancy.

In most cases, the HSG was performed using a cervical vacuum cup. More than 90 percent of women in both the intervention and control group used pain medication prior to the HSG. HSG procedure characteristics and outcomes are reported in Table 2.

Table 2. HSG procedure characteristics and outcomes.

Characteristic	Intervention (n=69)	Control (n=65)
HSG Completed (%)	68 (98.6)	62 (95.4)
Instrumentation (%)		
Cervical vacuum cup	42 (60.0)	41 (63.1)
Hysterophore	5 (7.2)	4 (6.1)
Ballooncatheter	18 (26.0)	18 (27.7)
Multiple	4 (5.8)	2 (3.1)
Tenaculum (%)	11 (15.9)	9 (13.8)
Bilateral tubal patency (%)	48 (69.6)	53 (82.8)
Type of contrast (%)		
Oil-based	66 (95.7)	61 (93.8)
Water-based	3 (4.3)	4 (6.2)
Amount of contrast (ml) (mean, SD) #	9.4 (3.44)	8.6 (3.43)
Pain medication used (%)	65 (94.2)	62 (95.4)
Type of pain mediation used (%) ^		
Paracetamol/panadol	18 (36.5)	12 (18.5)
NSAID	54 (78.3)	56 (86.2)
Opioids	0	0
Other	0	1 (1.5)
Mean Anxiety score (APAIS) (mean, SD) *	9.2 (2.29)	10.1 (2.38)

#reported for n=65 in the intervention group and n=61 in the control group. ^Some patients used multiple types of pain medication. * measured on a scale of 3 – 15.

Primary outcome

The overall mean pain score was 5.0 (SD 2.10) in the VR during HSG group, compared to 4.9 (SD 2.13) in the control group (p=0.915). Women reported a mean peak pain score of 6.8 (SD 2.25) using VR during HSG, compared to a peak pain score of 6.6 (SD 2.40) in the control group (p=0.574). The mean proportion of time women thought about being in pain during HSG, measured on the 5-point Likert scale (scale 1-5), was not significantly different between the intervention and control group (3.3 (SD 0.99) versus 3.5 (SD 1.10), p=0.159) (Table 3).

Table 3. Primary outcome.

Outcome	Intervention (n=69)	Control (n=65)	p-value^
Overal pain (VAS) (mean (SD))[1]	5.0 (2.10)	4.9 (2.13)	0.915
Peak pain (VAS) (mean (SD))[1]	6.8 (2.25)	6.6 (2.40)	0.574
Time spend thinking about pain (mean (SD)*	3.3 (0.99)	3.5 (1.10)	0.159
Pain medication used within 24h (%) #			0.360
Yes	9 (18.8)	10 (21.3)	
No	37 (77)	37 (78.7)	
Unknown	2 (4.2)	0	
Type of pain medication used (%) $			
Paracetamol	3 (33.3)	6 (60)	0.317
NSAIDs	8 (88.9)	6 (60)	0.779

Pain scores (in VAS) immediately after HSG, measured in peak pain and overall experienced pain. ^statistics with use of Fisher's exact test or Chi-square test *Reported for n=68 intervention group, n=63 control group, measured using a 5-point Likert scale. # Reported for n=48 intervention group, n=47 control group. $ Reported for n=65 intervention group, n=62 control group. [1]measured using a VAS-ruler, scale 0.0 – 10.0cm

To explore whether other factors might have influenced both peak and overall pain scores, multiple regression was performed that included anxiety levels prior to HSG, tubal patency, VR expectations, previous experience with VR, the type of instruments used (i.e. cervical vacuum cup, a hysterophore or a balloon catheter) and the type of contrast used (i.e. water-based contrast, oil-based contrast). The results of these multiple regression models are presented in Supplementary Table 1. Overall pain scores only seemed to be influenced by the expectation that VR would be a good distraction prior to the HSG procedure, with a standardized coefficient of 2.995 (p=0.005). Patients who expected a greater distraction by VR, reported a higher overall pain score. No other factors were found to have influenced overall pain scores.

Secondary outcomes: Satisfaction, Virtual Reality expectations and experience

Satisfaction

Women in the intervention group reported a mean satisfaction score on the 5-point Likert Scale (scale 1-5) of 4.07 (SD 0.84), whereas women in the control group reported a mean satisfaction score of 3.74 (SD 1.02) (p-value 0.197). While women in the intervention group reported similar willingness to undergo another HSG with VR, compared to the control group (3.43 (SD 1.35) versus 3.61 (SD 1.37) respectively, p-value 0.608), women in the intervention group reported significant less willingness to undergo another HSG without VR than the control group (2.72 (SD 1.25) versus 3.63 (SD 1.38) respectively, p-value 0.005) (Table 4).

Table 4. Patient satisfaction.

Outcome	Intervention n=26	Control n=23	p-value
Satisfaction score (mean (SD))[1]	4.12 (0.86)	3.78 (0.90)	0.197
Another HSG with VR (mean (SD))[1]	3.43 (1.35)	3.61 (1.37)	0.608
Another HSG without VR (mean (SD))[1]	2.73 (1.25)	3.63 (1.38)	0.005
Advise an HSG with VR to others (mean (SD))[12]	4.23 (0.91)	N/A	

Full questions (range 1-5): 'How satisfied are you with the HSG with/without VR?' 'Would you be willing to undergo another HSG without VR?' 'Would you be willing to undergo another HSG with VR?' 'Would you advise to use VR if someone else had to undergo HSG?'

[1] measured using a 5-point Likert scale, scale 1-5.

[2] Only answered by women in the intervention group.

Virtual Reality expectations

A subset of women (n=50) was asked about their expectations of VR prior to group allocation, using an additional questionnaire with multiple choice questions. Most women expected that VR would be a good distraction from the HSG (78%), while a considerably smaller portion of women (18%) expected VR to reduce their discomfort. Prior to the procedure, 52% of women reported they preferred relaxation exercises, 38% preferred simple movies, while 10% had no preference. Real life movies were preferred over animated movies by 67% (see Supplementary Figure 1a-c).

Virtual Reality experience

The predefined possible side-effects of Virtual Reality (nausea, dizziness, headache and blurred vision) did not occur more frequent in the intervention group than in the control group. Eight patients removed the VR headset during the procedure because of discomfort (Supplementary Table 2).

In the subset of women (n=50) that were asked on their expectations of VR prior to group allocation, some additional, more in-depth questions were asked after the HSG procedure to women in the intervention group (n=26) regarding their VR experiences after they underwent the HSG procedure. All questions used a 5-Point Likert scale ranging from 1 – 5, where 1 represented 'not at all' and 5 represented 'completely/all the time'. Women reported that they were able to concentrate on the VR application between some and most of the time with a mean score 3.19 (SD 0.94). They were distracted by VR between some and most of the time, with a mean score of 3.54 (SD 0.76). Women reported to miss sounds from the VR headset between a little and some of the time, with a mean score of 2.50 (SD 1.30).

Over sixty percent (16/26) of women in the subset that had received VR reported that their VR expectations were met. Almost forty percent (10/26) reported that their expectations

were only partially, or not met (Table 5). When asked an open-ended question why their expectations were not (fully) met, women stated that pain predominated the VR application, or that the installed video applications were insufficiently engaging to distract them completely. When asked what other VR applications women would have liked to see, they noted that they would have liked to watch an episode of a comedy series, other nature sceneries or short films with baby animals (kittens/puppies).

Table 5. VR experiences.

	Total n=26*
Distracted by VR (mean, SD)	3.54 (0.76)
Able to concentrate (mean, SD)	3.19 (0.94)
Feeling into the VR environment (mean, SD)	2.90 (1.14)
Missing sound (mean, SD)	2.50 (1.30)
Satisfaction with film options (%)	14 (53.8)
VR expectations were met (%)	
Yes	15 (57.7)
No	1 (3.8)
Partially	10 (38.5)

Full questions (range 1-5, 'not at all' to 'completely/all the time'): 'To what extend did the movie(s) distract you from the HSG?' 'To what extend were you able to focus on the movie(s)?' 'There was no sound with the movie(s). Did you miss this?'

* These questions were answered by women from the subset (n=50) that had received the VR intervention (n=26).

DISCUSSION

Summary of findings

In this RCT the use of Virtual Reality (VR) to reduce pain during HSG was not effective. Patients in both the intervention group and control group reported similar peak pain scores and overall pain scores. In addition, no significant differences were seen in satisfaction score and amount of time spend thinking of their discomfort.

Comparison to other studies

In a recent RCT of 82 women, VR was applied during an operative hysteroscopy (19). Women either received VR during hysteroscopy in addition to standard care, or standard care alone. Virtual Reality did not result in significant lower pain scores during the procedure (median pain score 5.0 in both the intervention and control group, p=0.67). In contrast, an RCT of 40 women that applied VR during hysteroscopy, found a statistically significant decrease in both worst pain and average pain (mean reduction 2.2 and on a

10-point VAS). Additionally, they reported a statistically significant reduction in mean anxiety score when VR was used during office hysteroscopy (15).

The mean duration of the hysteroscopy procedure in the study by Deo et al. was less than 4 minutes in both the intervention and control group. Although we did not record procedure length in our study, the average time of a HSG procedure is expected to be longer. The applications on our VR headset had a duration of approximately two minutes each. Women subsequently had to choose another VR application or were transitioned into a new application automatically, which might have reduced the submersion in the VR environment and therefore negatively affect procedural distraction and pain reduction.

A recent study provided VR during postpartum episiotomy repair, adding it to standard care consisting of lidocaine infiltration (5ml of 2% solution). Virtual Reality was found to decrease pain significantly during episiotomy repair when compared to standard analgesics alone (VAS 9.0±12.6 versus 23.6±19.8 respectively with an intergroup p-value of 0.038, scale 0.0-100.0) (16). It should be noted that in this small study (n=30), the mean episiotomy depth was shorter in the intervention group (p=0.042), possibly resulting in a shorter procedure time. This could possibly influence the main findings.

A group in China is currently conducting a similar study to ours, randomizing 200 women between an HSG with VR or a blank control during HSG (6) . In addition to pain, anxiety, satisfaction, acceptance and immersion, they will also record physiological parameters (i.e. pulse rate, blood pressure and oxygen saturation) to relate to the primary outcome. Participants in the intervention group will receive a VR system that includes noise-reducing headphones to convey sound, in contrast to our VR system without auditory stimuli.

Explanations

The lack of effect in overall pain and peak pain score during HSG using VR versus no VR can have several reasons. One explanation might be that because of the nature of this study, there was some emphasis on the fact that the HSG procedure could be painful. During study counselling, it was mentioned to women that the HSG procedure might be somewhat uncomfortable, similar to heavy menstrual cramping. Therefore, women might have expected the HSG to be painful, prior to the procedure. This might be an explanation for the similar reported pain scores after the HSG procedure in the intervention- and control group. Women randomized for VR glasses during HSG might not have experienced as much pain relief from the intervention as they expected, whereas women randomized for HSG without VR glasses reported that the procedure was “not as uncomfortable as they expected”. It might also explain why expectation prior to the HSG procedure influenced the reported pain scores. A previous study found that participants experienced more pain

when they had received high-pain cues, compared to participants receiving low-pain cues (20). However, the mean overall pain scores found in our study (VAS 5.0 in the intervention group and 4.9 in the control group) are in line with pain scores found as a secondary outcome in a previous study (VAS 5.0 median) (9). This secondary analysis additionally found that women receiving oil-based contrast had a higher ongoing pregnancy rate if they experienced moderate to severe pain during HSG than when they experienced mild pain (9). This difference was not seen in women who received water-based contrast during HSG. The hypothesis for this pregnancy enhancing effect is that oil-based contrast has a greater potential to flush away pregnancy hindering debris because of its chemical or physical characteristics. The contrast fluid is thought to create a pressure build-up behind the debris, causing pain, until the debris comes loose and the pressure releases. More pain anticipated prior to the procedure either might have influenced both the intervention and the control group, or might not have influenced the pain scores at all. In our study, we did not include ongoing pregnancy rate as a secondary outcome. Therefore, we could not compare the reported pain scores to pregnancy rates.

It is known that tubal blockage causes a higher level of pain during HSG than bilateral tubal patency (21). In our study, we saw a non-significant difference in the prevalence of bilateral tubal patency between the intervention- and the control group (75.4% versus 80.0%, respectively. P=0.481). Bilateral tubal blockage was diagnosed in only one woman in the control group and two women in the intervention group. In addition, we found that tubal patency did not significantly influence the reported (average) pain scores. This is contrary to the findings by Szymusik et al. where a significant difference was reported in VAS scores during contrast instillation by women with any abnormality in tubal patency, compared to women with bilateral tubal patency (VAS-score of 7.30 in women with abnormal tubal patency versus 5.80 in women with bilateral tubal patency, $p<0.001$) (21).

For our study, we chose to turn off the sound of the VR application. We deemed it important that the caregiver was able to communicate with the patient and explain different steps of the HSG procedure. This might have influenced the level of immersion of the patient, and thereby the pain reduction properties of VR. Previous research demonstrated that the highest pain tolerance was reported by patients exposed to both visual VR and its supplementary sound, compared to visual VR alone (22). Won et al. stated the five criteria for successful VR in their paediatric study (23). These five criteria comprised presence (subjective feeling that the user is inside the VR environment), interactivity (involvement of the user into the VR environment, e.g. by incorporating tracked body movements), social interactions (a VR environment can offer an alternative platform to build relationships), customization (adjusting the VR environment to the patient preferences) and embodiment (user is embodied in the VR environment through

their avatar by using body trackers). Relevant for our study were the following criteria: presence, interactivity and customization. By optimizing the participants' presence, or immersiveness, in the VR environment, they might report lower overall and peak pain scores. Interactivity of the VR program is related to presence as well. Both contribute to pain reduction: a higher level of interaction with the VR environment and presence in the VR environment, might lead to better distraction from the procedure and therefore a higher pain threshold and less experienced pain. Finally, customization is an important factor in the effectiveness of Virtual Reality. Involvement in the choice or even design of the VR application can contribute to pain reduction. In our study, women were able to choose between the different VR applications, i.e. nature films and breathing exercises. However, women were not able to design or modify their VR application.

In obstetrics, it is known that one-to-one coaching of labouring women reduces the level of experienced pain and the need for pharmacological pain relief (24). In daily practice, coaching with use of breathing techniques is often used during HSG if the physician or radiology technician feels this is beneficial for the woman. With our intervention, VR during HSG, it is often not possible to coach the woman at the same time because it might decrease immersiveness. It is therefore possible that our results do not show that VR is not effective, but rather show that VR is not more effective in reducing discomfort as compared to coaching with breathing techniques. However, we feel that a comparison of regular care without coaching as a control group would not have been ethical.

Implications

As mentioned before, the VR applications that we used can potentially be altered to provide a more engaging and immersive experience. However, after our experience counseling potential participants, we feel that for some women a maximum immersive experience would be optimal while other women want to have the ability to distract themselves when desired but be present during the procedure at other times. Therefore, a more personalized VR application might potentially lead to lower pain scores. More information on what type of VR suits which patient characteristic is necessary. Patient characteristics can include preferred coping strategies, preferred level of control during a medical procedure, preferences on the types of movie content (comedy, nature, animals) and/or whether they would prefer to play games. For future studies, it could be advised to further optimize this customization by studying the patient characteristics (i.e. preferences on coping mechanisms and level of control, type of VR application), presence and interactivity. This could contribute to VR immersiveness, thereby optimizing pain reduction during the HSG procedure when applying a VR headset. We would therefore recommend future studies to pay attention to VR immersiveness, and to optimize VR immersiveness to hopefully improve VR effectivity.

Strengths and limitations

Our study is, to our knowledge, the first to report on the efficacy of VR as additional pain relief during HSG. In addition, our study was performed in a multicentre setting. Both locations of our university hospital, as well as a non-academic teaching centre, have recruited participants, resulting in a varied study population. The questions on expectations of the use of VR prior to the procedure and the experience of patients in the intervention group afterwards enabled us to get a better understanding of the expectations and preferences women have.

Our study also has weaknesses. A significant number of eligible women (50 of the 211 screened women) declined participation in this study, which might have generated selection bias. Most women who disclosed a reason for not wanting to participate in the study, mentioned that they preferred to have more control over the situation and would like to be able to see the physician perform the HSG. Others said they were already nervous to undergo the procedure and did not want to add to that with additional questionnaires. On the other hand, some women that were nervous to undergo the HSG procedure were more inclined to participate in the study, in the hopes of reducing their discomfort. It is possible that anxiety resulting in renunciation of study participation, has led to the selection of a more homogenous study group in terms of anxiety prior to the procedure. However, anxiety scores were relatively high in both the intervention group and control group (9.2 and 10.1 respectively) which is consistent with literature (25). In addition, since women were also more inclined to participate in the study because of their anxiety, and this study was randomized, we expect that this effect is small.

Another weakness of our study is the relatively low level of VR immersion reported by women in the intervention group. On a scale of 1-5, the mean self-reported distraction posed by VR was rated 3.54 (SD 0.76), the mean ability to focus on VR was 3.19 (SD 0.93) and the lack of sound was rated 2.50 (SD 1.30). As women had to lay still in supine position during HSG, VR games requiring physical movement were disabled. Women were therefore able to choose from nature movies and relaxation exercises. This might have lowered the level interactivity of VR, while interactivity is one of the factors known to stimulate immersion and attentional involvement (26, 27). Stimulation of multiple senses at once is also thought to increase attentional involvement, since in our study sound was enabled to promote communication between the physician and woman this might further decrease immersion into VR. It might also be that, because both groups used analgesics the night before and 2 hours prior to the HSG procedure, the effect of VR on pain reduction was too small to exceed the effect of the analgesics.

In conclusion, Virtual Reality does not reduce pain experience during HSG when added to regular self-administered analgesics. Women with higher expectations of the effectiveness of VR prior to the procedure, reported higher pain levels afterwards.

REFERENCES

1. Aarts JW, van Empel IW, Boivin J, Nelen WL, Kremer JA, Verhaak CM. Relationship between quality of life and distress in infertility: a validation study of the Dutch FertiQoL. Hum Reprod. 2011;26(5):1112-8.
2. van der Linden MW, Westert GP, de Bakker DH, Schellevis FG. Tweede Nationale Studie naar ziekten en verrichtingen in de huisartsenpraktijk. Klachten en aandoeningen in de bevolking en in de huisartsenpraktijk. Utrecht/Bilthoven: NIVEL/RIVM; 2004.
3. Fertility problems: assessment and treatment [Internet]. 2013 [cited 17-02-2022]. Available from: https://www.nice.org.uk/guidance/cg156.
4. Bhattacharya S, Porter M, Amalraj E, Templeton A, Hamilton M, Lee AJ, et al. The epidemiology of infertility in the North East of Scotland. Hum Reprod. 2009;24(12):3096-107.
5. Dreyer K, van Rijswijk J, Mijatovic V, Goddijn M, Verhoeve HR, van Rooij IAJ, et al. Oil-Based or Water-Based Contrast for Hysterosalpingography in Infertile Women. N Engl J Med. 2017;376(21):2043-52.
6. Wang YL, Gao HX, Wang JS, Wang JH, Bo L, Zhang TT, et al. Immersive virtual reality as analgesia for women during hysterosalpingography: study protocol for a randomized controlled trial. Trials. 2020;21(1):102.
7. van Rijswijk J, van Welie N, Dreyer K, Pham CT, Verhoeve HR, Hoek A, et al. Tubal flushing with oil-based or water-based contrast at hysterosalpingography for infertility: long-term reproductive outcomes of a randomized trial. Fertil Steril. 2020;114(1):155-62.
8. Dreyer K, Out R, Hompes PG, Mijatovic V. Hysterosalpingo-foam sonography, a less painful procedure for tubal patency testing during fertility workup compared with (serial) hysterosalpingography: a randomized controlled trial. Fertil Steril. 2014;102(3):821-5.
9. van Welie N, Dreyer K, van Rijswijk J, Verhoeve HR, Goddijn M, Nap AW, et al. Treatment effect of oil-based contrast is related to experienced pain at HSG: a post-hoc analysis of the randomised H2Oil study. Hum Reprod. 2019;34(12):2391-8.
10. Serrano González L, Pérez-Medina T, Bueno Olalla B, Royuela A, de Los Reyes De La Cuesta M, Saéz de la Mata D, et al. Is hysterosalpingo-foam sonography (HyFoSy) more tolerable in terms of pain and axiety than hysterosalpingography (HSG)? A prospective real-world setting multicentre study. BMC Women's Health. 2022;22(41).
11. Hindocha A, Beere L, O'Flynn H, Watson A, Ahmad G. Pain relief in hysterosalpingography. Cochrane Database Syst Rev. 2015(9):CD006106.
12. Moore DE. Pain associated with hysterosalpingography: Ethiodol versus Salpix media. Fertil Steril. 1982;38:629-31.
13. Hoffman HG, Chambers GT, Meyer WJ, 3rd, Arceneaux LL, Russell WJ, Seibel EJ, et al. Virtual reality as an adjunctive non-pharmacologic analgesic for acute burn pain during medical procedures. Ann Behav Med. 2011;41(2):183-91.
14. Mallari B, Spaeth EK, Goh H, Boyd BS. Virtual reality as an analgesic for acute and chronic pain in adults: a systematic review and meta-analysis . Journal of Pain Research. 2019;Volume 12:2053-85.
15. Deo N, Khan KS, Mak J, Allotey J, Gonzalez Carreras FJ, Fusari G, et al. Virtual reality for acute pain in outpatient hysteroscopy: a randomised controlled trial. BJOG. 2021;128(1):87-95.
16. JahaniShoorab N, Ebrahimzadeh Zagami S, Nahvi A, Mazluom SR, Golmakani N, Talebi M, et al. The effect of virtual reality on pain in primiparity women during episiotomy repair: a randomize clinical trial. Iran J Med Sci. 2015;40:219-24.
17. Mosso Vázquez JL, Wiederhold BK, Miller I, Wiederhold MD. Virtual Reality assisted anesthesia during upper gastrointestinal endoscopy: report of 115 cases. EMJ Innov. 2017;1(1):75-82.

18. Bekelis K, Calnan D, Simmons N, MacKenzie TA, Kakoulides G. Effect of an Immersive Preoperative Virtual Reality Experience on Patient Reported Outcomes: A Randomized Controlled Trial. Ann Surg. 2017;265(6):1068-73.
19. Fouks Y, Kern G, Cohen A, Reicher L, Shapira Z, Many A, et al. A virtual reality system for pain and anxiety management during outpatient hysteroscopy-A randomized control trial. Eur J Pain. 2022;26(3):600-9.
20. Jepma M, Koban L, van Doorn J, Jones M, Wager TD. Behavioural and neural evidence for self-reinforcing expectancy effects on pain. Nature Human Behaviour. 2018;2(11):838-55.
21. Szymusik I, Grzechocinska B, Marianowski P, Kaczynski B and Wielgos M. Factors influencing the severity of pain during hysterosalpingography. *Int J Gynaecol Obstet* 2015;**2:**118-122
22. Johnson S, Coxon M. Sound can enhance the analgesic effect of virtual reality. R Soc Open Sci. 2016;3(3):150567.
23. Won AS, Bailey J, Bailenson J, Tataru C, Yoon IA, Golianu B. Immersive Virtual Reality for Pediatric Pain. Children (Basel). 2017;4(7).
24. Bohren MA, Hofmeyr GJ, Sakala C, Fukuzawa RK, Cuthbert A. Continuous support for women during childbirth. Cochrane Database Syst Rev. 2017;7:CD003766.
25. La Fianza A, Dellafiore C, Travaini D, Broglia D, Gambini F, Scudeller L, et al. Effectiveness of a single education and counseling intervention in reducing anxiety in women undergoing hysterosalpingography: a randomized controlled trial. ScientificWorldJournal. 2014;2014:598293.
26. Di Guardo F, Palumbo M. Immersive virtual reality as tool to reduce anxiety during embryo transfer. J Obstet Gynaecol. 2022;42(5):802-8.
27. Slater M, Lotto B, Arnold MM, Sanchez-Vives MV. How we experience immersive virtual environments: the concept of presence and its measurement. annuario de psicologia. 2009;40:193-210.

SUPPLEMENTARY MATERIAL

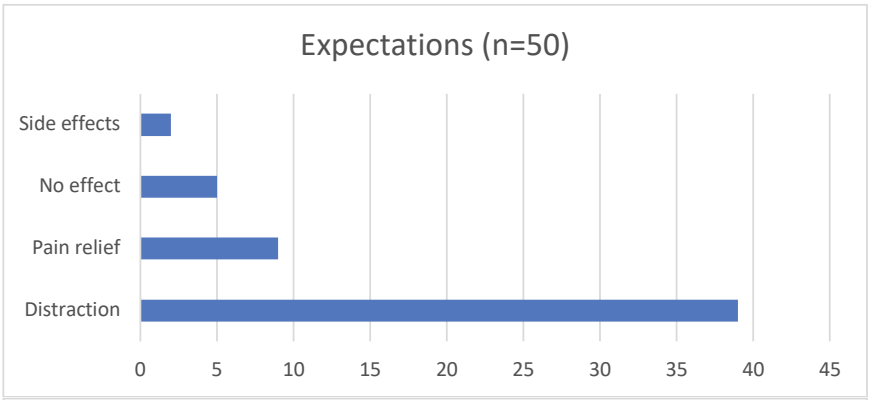
Supplementary Table 1. Regression model

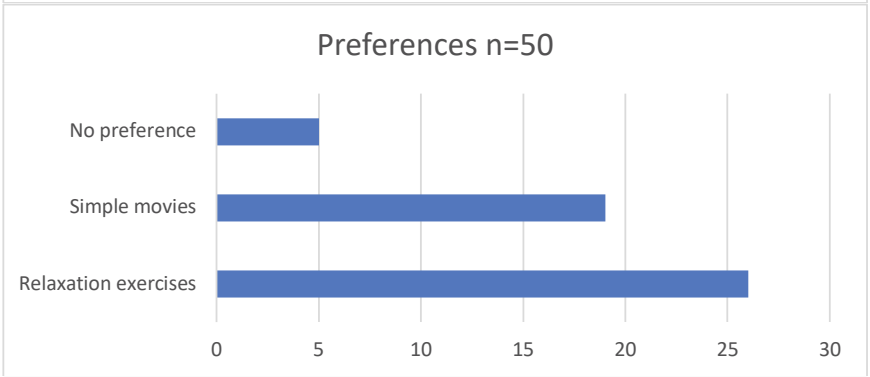
	B (standard error)	Beta (t value)	p-value
APAIS total	0.205 (0.135)	0.228 (1.523)	0.136
Expect distraction	2.704 (0.973)	0.572 (2.778)	0.009
Expect pain reduction	-0.330 (0.789)	-0.065 (-0.418)	0.678
Expect no effect	1.056 (1.272)	0.162 (0.830)	0.412
Experience with VR	0.211 (0.546)	0.053 (0.387)	0.701
Semm cup	-0.033 (0.982)	-0.008 (-0.033)	0.974
Schultze	0.950 (1.576)	0.194 (0.603)	0.550
Balloon catheter	0.585 (1.055)	0.124 (0.555)	0.582
Tenaculum	2.096 (1.373)	0.469 (1.526)	0.135
Type of contrast	0.534 (0.833)	0.095 (0.641)	0.525
Randomization group	0.673 (0.578)	0.170 (1.165)	0.252

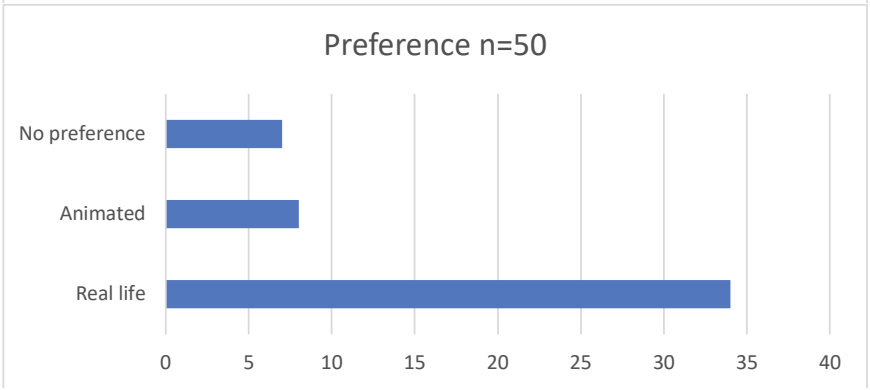
Supplementary Table 2. Side effects

Side effects during HSG (mean, SD)		Intervention n=68	Control n=64	p-value
	Nausea	1.6 (1.08)	1.7 (1.18)	0.422
	Dizziness	1.6 (0.98)	1.7 (1.08)	0.583
	Headache	1.1 (0.57)	1.2 (0.48)	0.482
	Blurred vision	1.3 (0.84)	1.2 (0.41)	0.489
Removal of VR headset		7	N/A	

Supplementary Figure 1a, b & c.







CHAPTER 7

Visual tubal patency tests for tubal occlusion and hydrosalpinx

Rachel Tros, Kimmy Rosielle, Carolien Koks, Velja Mijatovic, Marlies Y Bongers, Ben Willem J Mol, Rui Wang

Cochrane library. Version published: 09 June 2022. https://doi.org/10.1002/14651858.CD014968

ABSTRACT

Objectives

This is a protocol for a Cochrane Review (diagnostic). The objectives are as follows:

To determine and compare the diagnostic accuracy of visual tubal patency tests (hysterosalpingography (HSG), sono-hysterosalpingography (sono-HSG), magnetic resonance hysterosalpingography (MR-HSG), and outpatient transvaginal hydrolaparoscopy (THL)) for the diagnosis of tubal occlusion.

Secondary objectives

To determine and compare the diagnostic accuracy of visual tubal patency tests (HSG, sono-HSG, MR-HSG, and outpatient THL) for the diagnosis of hydrosalpinx.

To evaluate heterogeneity with regards to population characteristics (population risk stratification) and index test characteristics (contrast media, technology, operator skills).

BACKGROUND

Infertility, defined as the failure to conceive within 12 months of regular unprotected sexual intercourse, occurs in at least 12% of the couples who wish to conceive (1, 2). Around 18% to 33% of couples with infertility present with tuboperitoneal pathologies such as blocked or damaged Fallopian tubes (3-5). As the Fallopian tubes are essential for transportation of the spermatozoa, the ovum and the embryo (6), bilateral occluded tubes exclude the chance of natural pregnancy. Therefore, bilateral tubal occlusion formed the basis of the development of in vitro fertilization (IVF) and was the earliest indication for IVF (7).

Most diagnostic protocols for fertility assessment include a test to rule out tubal occlusion (8, 9). During such tubal patency tests, a contrast agent is flushed into the uterus and through the Fallopian tubes, visualizing tubal patency. Diagnostic laparoscopy with methylene blue dye tubal patency testing, also known as chromopertubation, is generally accepted as the reference standard (9). However, due to its invasiveness and costs, alternative less invasive tests have been carried out as replacements. These visual tubal patency tests have evolved alongside the development of radiography, ultrasonography and laparoscopy, including hysterosalpingography (HSG), sono-hysterosalpingography (sono-HSG), magnetic resonance hysterosalpingography (MR-HSG), and outpatient transvaginal hydrolaparoscopy (THL). The choice of these visual tubal patency tests varies in different settings.

Visual tubal patency tests can be used to diagnose tubal, uterine and other pelvic conditions. The most important tubal conditions are bilateral tubal occlusion, unilateral tubal occlusion and hydrosalpinx. The diagnoses of these conditions will directly guide clinical management, so they will be the focus of this Cochrane Review.

Target condition being diagnosed

Target conditions of interest are tubal occlusion and hydrosalpinx.

1. Tubal occlusion: women with untreated bilateral occlusion have no chance of a natural pregnancy, as there is no way for the ovum and spermatozoa to meet, and these women can benefit from early IVF (7). Therefore, women diagnosed with bilateral tubal occlusion are mostly offered IVF directly, although IVF is not available to all couples worldwide due to differences in health care systems and reimbursements. IVF can be preceded by laparoscopic surgery to optimize pelvic anatomy. Management in women with unilateral tubal blockage is more diverse, as in these women the patent Fallopian tube still facilitates transport of the ovum,

spermatozoa and embryo. Mostly unilateral tubal patency is treated the same as bilateral tubal patency, as unilateral tubal patency does not reduce pregnancy outcomes significantly (10). There are studies reporting on lower odds of pregnancy when unilateral distal tubal occlusion is detected in comparison to proximal tubal occlusion (11). This observed difference between proximal and distal tubal occlusion may result from inherent diagnostic limitations of HSG or may reflect different underlying pathologies that differentially affect pregnancy outcomes. However, proximal or distal occlusion cannot be identified by all index tests, so we will not differentiate between proximal or distal occlusion in this review.

2. Hydrosalpinx: the other condition of interest is hydrosalpinx. It refers to the distension of the Fallopian tube due to distal tubal occlusion and fluid accumulation, and the most common cause is a previous episode of pelvic inflammatory disease (12). A hydrosalpinx has a negative impact on fertility outcomes through different mechanisms. Removal or ligation of the hydrosalpinx has a positive effect on clinical pregnancy rates before assisted reproductive technology (ART) (13).

Other conditions that are not the focus of this review, but can be detected during visual tubal patency tests, are endometriosis (which can be visualized during THL), peritoneal/pelvic adhesions (sono-HSG and THL) and intrauterine pathology (HSG and sono-HSG). Endometriosis is seen in about 25% to 40% of women with infertility (14, 15). Pelvic adhesions, caused by previous surgery, pelvic inflammatory disease or endometriosis, may interfere with ovum pickup if they are distorting the anatomy of the ovary and Fallopian tube. Intrauterine pathology as myomas, polyps or intrauterine adhesions, as well as congenital uterine anomalies, can be detected by some of the visual tubal tests. These intracavitary conditions might all have some effect on fertility outcomes (8, 16).

Index test(s)

We will consider the following four main groups of index tests.

1. Hysterosalpingography (HSG): this uses serial X-ray or fluoroscopy images during injection of an iodine-containing contrast medium through the cervical canal into the uterus and subsequently the Fallopian tubes. Different instruments, such as a reusable metal cannula (hysterophore or Jarcho cannula), a 5-French balloon catheter or a (modified) cervical vacuum cup device, as well as different iodine-containing contrast media, oil-based or water-based, can be used. HSG is contraindicated in women with an allergy to iodine-containing contrast media. HSG is a safe and widely accepted procedure in the outpatient setting, but it needs to be performed in a radiology department. HSG is well-tolerated, although more painful

than sono-HSG (17) or THL (18). In addition to its advantage of evaluating the uterine cavity and tubal patency, it has a potential therapeutic effect when an oil-soluble contrast medium is used, with a higher chance of clinical pregnancy and live birth rates (19). Choice of contrast medium, operator skill and the observer interpreting the HSG are likely to be potential sources of heterogeneity (20).

2. Sono-hysterosalpingography (sono-HSG): this includes both hysterosalpingo-foam sonography (HyFoSy) and hysterosalpingo-contrast sonography (HyCoSy). Overall, this test is based on ultrasound, in which an echogenic medium is used to assess the uterine cavity and tubal patency. Many different sono-HSG techniques are performed, with differences in two- or three-dimensional ultrasound modality; vaginal or abdominal ultrasound; contrast type (commercially available foam as well as normal saline, saline and air or galactose, or combinations of these); or the usage of colour doppler sonography (21). The advantages of these tests are that they can be performed in an outpatient setting without a radiology department (offering the possibility of a one-stop fertility evaluation), and are generally well tolerated (17). Furthermore, when compared to HSG, the procedure does not require exposure to radiation or iodine-containing contrast media (22). In addition to tubal patency, the uterine cavity and myometrium, as well as both ovaries, can be assessed during the procedure (23). It is likely that choice of contrast, operator skill and test technology influence the diagnostic quality.

3. Magnetic resonance hysterosalpingography (MR-HSG): this is similar to HSG. It uses MR-imaging instead of X-ray or fluoroscopy, and the contrast medium is a gadolinium-based solution, available from different manufactures and prepared in different ways (24). Similar to HSG, the procedure can be performed in an outpatient setting when a radiology department is available, and it is well tolerated (25, 26). It also avoids exposure to radiation and iodine-containing contrast media, and can be used to diagnose (deeply infiltrating) endometriosis, uterine and ovarian anomalies. In comparison to sono-HSG, the advantage of MR-HSG is that is not operator dependent, with a better reproducibility (24, 26).

4. Transvaginal hydrolaparoscopy (THL): also known as transvaginal endoscopy or fertiloscopy, this technique uses hydroflotation for exploration of the pelvic cavity. A small diameter optic is inserted transvaginally through an incision in the vaginal posterior fornix, after the pelvis is filled with warm normal saline for pelvic cavity distention. By using a dye, mostly methylene blue, tubal patency can be tested (27). Different instruments, disposable or reusable, can be used for THL (28). THL is a known, safe and well-tolerated procedure, which can be performed in an

outpatient setting under local anaesthesia (28, 29, 18)). An advantage is the direct visualization of the female genital tract, thus allowing the evaluation of hydrosalpinx, endometriosis, and pelvic adhesions next to the tubal blockage. It is possible that the experience of the operator influences the success rate of THL.

Clinical pathway

There has been a wide range of variation in visual tubal patency tests during fertility workup, at both national and international levels (8, 9). In general, a comprehensive medical history is obtained as the first step to explore the possible causes of female-factor infertility. Next, physical examination and transvaginal ultrasound assessment are performed. In some settings, tubal patency is always then tested (8, 9), while in other settings, tubal testing is considered based on findings from medical history, physical examination and serological testing (chlamydia antibody testing; CAT) and only women with a high risk for tubal pathology will undergo tubal testing (30). Women are usually considered as having a high risk for tubal pathology when they have had a history of chlamydia infection or a positive CAT, pelvic inflammatory disease or peritonitis, or when they have been diagnosed with endometriosis or have had pelvic surgery in the past (31, 32).

A visual tubal patency test can be used as a triage or as a replacement test. When used as a triage test, women will undergo laparoscopy only when occlusion is suspected or the visual tubal patency test shows indeterminate findings. However, the aim of laparoscopy in current practice is more often to select women who may benefit from therapeutic laparoscopy, rather than to select women for diagnostic laparoscopy. For example, in the Federation Medical Specialists (FMS) guideline (30), visual tubal patency tests are performed in high risk women as a triage test to select women who require laparoscopy. Laparoscopy without prior visual tubal patency testing is reserved only for those women with severe endometriosis or hydrosalpinges, where the diagnostic procedure and therapeutic laparoscopy are combined at the same time. In most other settings a visual tubal patency test is used as a replacement for the reference standard. The outcome of this test will then be used to determine if fertility treatment is necessary. Fertility treatment can be therapeutic laparoscopy or assisted reproduction, depending on the availability and preferences of the doctor and person undergoing treatment. An example is the NICE guideline (9), in which women with low risk for tubal pathology are offered a visual tubal patency test and those with high risk are offered a laparoscopy. Depending on the results, women with tubal obstruction can be offered tubal surgery, when appropriate expertise is available, or assisted reproduction directly.

The choice of visual tubal patency tests also varies in different settings, depending on the preference and skills of the clinician, the preference of the couples with infertility and the availability of tubal testing methods in the clinic. In different geographical and economic contexts, costs, availability and the accessibility of these testing methods will differ. However, if available and accessible, this protocol hopes to answer the question of which visual tubal patency test should be advised above others as a replacement test for laparoscopy to diagnose tubal patency.

Alternative test(s)

Alternative tests are not applicable, as all visual tubal patency tests will be reviewed in this protocol.

Rationale

Over the last two decades, new tubal patency tests (e.g. MR-HSG) have been emerging, as well as new contrast media or test technology for existing tubal patency tests. Therefore, it is important to summarise all the evidence on the accuracy of individual tests, and to compare different tests' accuracies. However, there is no Cochrane Review on this topic. As visual tubal patency tests are all less invasive than diagnostic laparoscopy and are well-tolerated in an outpatient setting, it seems fair to offer such a test instead of the reference standard diagnostic laparoscopy. Nevertheless, the diagnostic accuracy of these tests is less acknowledged in clinical decision-making about the choice of tubal testing method. Currently, there is no consensus in terms of how different types of visual tubal patency test compare to each other. Before replacing the reference standard, it is important to understand the diagnostic accuracy of each individual visual patency test and to compare the diagnostic accuracy when possible. Couples with infertility will benefit from this research as it will guide clinicians to select the most suitable visual tubal patency test for the individual couple.

OBJECTIVES

To determine and compare the diagnostic accuracy of visual tubal patency tests (hysterosalpingography (HSG), sono-hysterosalpingography (sono-HSG), magnetic resonance hysterosalpingography (MR-HSG), and outpatient transvaginal hydrolaparoscopy (THL)) for the diagnosis of tubal occlusion.

Secondary objectives

To determine and compare the diagnostic accuracy of visual tubal patency tests (HSG, sono-HSG, MR-HSG, and outpatient THL) for the diagnosis of hydrosalpinx.

To evaluate heterogeneity with regards to population characteristics (population risk stratification) and index test characteristics (contrast media, technology, operator skills).

METHODS

Criteria for considering studies for this review

Types of studies

We will include studies on the diagnostic test accuracy of a single index test and studies on the comparative diagnostic test accuracy of two or more index tests.

For the diagnostic test accuracy of a single index test, we will include single-gate studies, in which one of the index tests (defined below) is compared with the reference standard within a timeframe in which the tubal status is unlikely to be changed (within three months).

For the comparative diagnostic test accuracy of two or more index tests, we will include the two following types of studies.

1. Studies with fully-paired direct comparisons. In these studies, the participants receive two or more index tests and the reference standard.
2. Randomized controlled trials that directly compares two or more index tests. In all arms, the index test should be followed by diagnostic laparoscopy as the reference standard.

We will exclude two-gate studies, as these study designs are likely to overestimate sensitivity and specificity. Furthermore, we will exclude studies with the primary endpoint of prognostic capacity for fertility outcomes, as well as diagnostic accuracy studies for sterilization purposes.

We will exclude studies with a sample size of fewer than 50 participants, given the relatively low prevalence of bilateral tubal occlusion across all risk groups. Although this threshold may be considered arbitrary, both sensitivity and specificity could be unreliable or biased in studies with smaller sample size.

Participants

We will include participants with infertility undergoing a visual tubal patency test and a diagnostic laparoscopy. We will include participants who have been trying to conceive for one year or more, both with low and high risk of tubal pathology, as well as unselected

participants (i.e. participants undergoing a visual tubal patency test without knowing about their risk of having tubal pathology). High risk incorporates all women with a positive history of pelvic inflammatory disease/chlamydia or who are CAT positive, those with extensive abdominal or tubal surgery in the past, and those with abnormalities like endometrioses/possible hydrosalpinx discovered during physical examination. Participants with low risk on tubal pathology are those with no previously-mentioned conditions for high risk. We will include participants who have had previous tubal testing only when the outcome of this test was not used to select the participants, as we will not include studies with a two-gate design. We will exclude participants undergoing tubal testing after refertilization (surgery to undo a tubal sterilization).

Index tests

We will include the following types of index texts.

- HSG, with either oil-based or water-based contrast.
- MR-HSG, including all techniques/MR-protocols or contrast media used.
- THL, transvaginal endoscopy or fertiloscopy, conducted with reusable instruments or disposable trocars.
- Sono-HSG (including HyFoSy and HyCoSy), used with commercially available foam, saline, saline and air or galactose, or combinations of these. We will exclude studies conducted with contrast that is no longer available (Echovist; galactose microparticles; Bayer Schering Pharma AG, Berlin, Germany). Furthermore, we will include studies that use two- or three-dimensional modality, with or without colour doppler.

Target conditions

We will consider tubal occlusion as a dichotomous diagnosis for all tests, i.e. occluded or patent (not occluded). As the unit of analysis will be at the individual level, due to its clinical importance, we will treat bilateral tubal occlusion and at least one-sided tubal occlusion as two separate conditions, instead of a threshold.

Similarly, we will also consider hydrosalpinx as a dichotomous diagnosis.

Reference standards

Laparoscopy with methylene blue dye tubal patency testing is the reference standard. All participants in the included studies should undergo this reference standard to avoid verification bias.

We will only include video-assisted laparoscopy, as this is less operator-dependent. We will exclude studies on direct visualization laparoscopy or those using CO2-pertubation or indigo carmine dye for tubal testing during laparoscopy, as well as the use of other dyes currently unknown to the authors. In addition, we will exclude studies with laparotomic tubal testing, and studies with another reference standard, for example using one of the index tests as the reference standard.

Search methods for identification of studies

Electronic searches

In collaboration with the Cochrane Gynaecology and Fertility (CGF) Group's Information Specialist, we will search the following electronic databases:

- CENTRAL (Cochrane Central Register of Controlled Trials) via the Cochrane Register of Studies Online (CRSO), web platform, to search from 1968 to present (Supplementary Table 1);
- MEDLINE, Ovid platform, to search from 1968 to present (Supplementary Table 2);
- Embase, Ovid platform, to search from 1980 to present (Supplementary Table 3);
- CINAHL (Cumulative Index to Nursing and Allied Health Literature; EBSCO platform, to search from 1968 to present (Supplementary Table 4).

For each database, we will use both index and free terms, and synonyms related to: infertility, tubal pathology, hysterosalpingography, hydrolaparoscopy, MR-HSG and sono-HSG.

We will also search trial registries for trials comparing two or more index tests, and for other eligible observational studies. We will search ClinicalTrials.gov (clinicaltrials.gov/), International Standard Randomised Controlled Trial Number (ISRCTN) registry (www.isrctn.com/), and the World Health Organization (WHO) International Clinical Trials Platform (ICTRP) Search portal (apps.who.int/trialsearch/).

As the reference test, laparoscopy with methylene blue dye tubal patency testing, was first reported in 1968 (33), we will use this as the earliest search date for those databases with an inception date prior to 1968.

Searching other resources

We will screen the reference lists of included studies and relevant systematic reviews for any additional trials. We will also search for ongoing and unpublished studies by approaching clinical experts and trialists in this field.

Data collection and analysis

Selection of studies

Two authors (RT and KR) will independently screen retrieved studies for eligibility on the basis of their titles and abstracts. If the study is potentially eligible, the same two authors will independently evaluate the full text for eligibility. A third author (CK or RW) will be involved to solve any disagreement at both stages. Where studies have multiple publications, we will collate multiple reports of the same study under a single study ID with multiple references. We will perform the study selection process in Covidence 2021 or other similar online platforms.

Data extraction and management

Two review authors (RT and KR) will perform the data extraction independently. When there is a disagreement between the two authors, a third author (CK or RW) will be consulted in the discussion. We will design a data extraction form for this review and pilot-test the form on three studies. We will collect the following data from the included studies: general information (first author, year of publication, country), participant characteristics (age, inclusion/exclusion criteria, numbers of participants, risk stratification (high/low risk for tubal pathology or unselected population), index test/reference standard details, two-by-two table for each outcome (true positives, true negatives, false positives, and false negatives), inconclusive tests and adverse events. When data for two-by-two tables are not available, we will calculate these data from the test accuracy results (sensitivity, specificity, positive predictive value and negative predictive value). Next, we will collect data on test-specific related conditions found. We will contact study investigators for information when needed.

Handling of inconclusive results

We will consider both valid inconclusive (intermediate or borderline) results and invalid inconclusive (indeterminate or uninterpretable) results in the analysis, as suggested by Shinkins (34). Participants with valid inconclusive results may receive further fertility treatment in clinical practice, but may also have another test in other settings. Therefore, we will treat all valid inconclusive results as positive (i.e. occluded) in the main analysis and as negative in a sensitivity analysis. Participants with invalid inconclusive results or procedure failures are more likely to have another test or a different index test in clinical practice, and some of these women may have conditions relevant to tubal pathology. Therefore, we will exclude invalid inconclusive results from the main analysis and treat them as positive (i.e. occluded) in a sensitivity analysis. We will evaluate the robustness of the findings by using different methods to handle the inconclusive results. Please refer to Sensitivity analyses.

Assessment of methodological quality

We will use the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) tool for the assessment of methodological quality of all included studies (35). We will evaluate the four domains in QUADAS-2 (patient selection, index test, reference standard, and flow and timing) for risk of bias, and the first three domains for concerns regarding applicability. For comparative diagnostic test accuracy studies, we will use the QUADAS-C tool (36) (Supplementary Table 5).

Two review authors (RT and KR) will evaluate the methodological quality of all included studies independently; we will resolve disagreements by involving a third author (CK or RW). We will pilot test both tools and only repeat this if it demonstrates problems during the first round.

Statistical analysis and data synthesis

Diagnostic test accuracy for each index test

We will perform the analysis for each index test separately. We will perform random-effects meta-analysis in a bivariate model (37). For each index test, we will present pairs of sensitivity and specificity with their 95% confidence intervals (CIs) for each study, as well as the pooled sensitivity and specificity, in a forest plot. We will then present the summary receiver operating characteristic (SROC) plot with summary points, and incorporate the seven domains of the QUADAS-2 tool into this SROC plot.

Unit of analysis

We will analyse the data on a participant level in the analysis, as this is more clinically relevant. However, if per participant data are not available and per tube is the unit of analysis in the majority of included studies, we will consider performing analysis on a per tube basis.

Comparative diagnostic test accuracy for different index tests

Direct comparison

We will include studies directly reporting two or more index tests compared with the reference standard in the primary analysis of comparative diagnostic test accuracy of different index tests. We will add a covariate for type of test in the bivariate model to compare the differences in test accuracy, and perform a likelihood ratio test to compare models (38). We will also present linked SROC plots, linking estimates of two different index tests from the same studies.

Indirect comparison

As indirect comparisons are prone to bias, we will only perform them as additional analyses if there are limited studies for direct comparisons. In this case, we will include studies that include one or more index test. We will evaluate the comparative diagnostic accuracy between HSG and other tests.

Procedure failure and adverse events

We will tabulate procedure failure and adverse events for all index tests and the reference test.

We will use Stata (38) and MetaDTA, an interactive online application for meta-analysis of DTA studies (39). When necessary, we will also use Review Manager 5.4.1 (40).

Investigations of heterogeneity

We will consider the following in the assessments of heterogeneity.

- Population characteristics: population risk stratification (high risk, low risk and unselected risk for tubal pathology).
- Index tests characteristics: HSG (oil versus water based contrast media; operator skills), sono-HSG (2D/3D versus 2D; different contrast media; use of colour doppler or not; operator skills), MR-HSG (different viscosity contrast media), THL (operator skills).

All these covariates are categorical variables. We will fit the models separately in different subgroups and perform visual inspections of SROC.

Sensitivity analyses

We will perform the following sensitivity analyses.

- Different approaches to handling inconclusive results:
  - o treating valid inconclusive results as negative (i.e. patent);
  - o treating invalid inconclusive results are positive (occluded).
- Limiting to studies at low risk of bias in the index tests and reference standard domains.

Assessment of reporting bias

We do not plan to evaluate reporting bias in this systematic review because statistical investigation of publication and reporting bias is not routinely recommended in DTA

systematic reviews, as stated in the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* (41, 42).

SUMMARY OF FINDINGS AND ASSESSMENT OF THE CERTAINTY OF THE EVIDENCE

For diagnostic accuracy of individual index tests, we will assess the certainty of the evidence according to the GRADE guidance 21 (43, 44). We will evaluate risk of bias, indirectness, inconsistency and imprecision, but will not assess publication bias for the reasons mentioned in Assessment of reporting bias. We will produce summary of findings tables for each index test, but these will be limited to one outcome only (bilateral tubal occlusion). We will present the number of studies/participants, study design, certainty assessment (risk of bias, indirectness, inconsistency and imprecision), summary of findings (numbers and 95% confidence intervals for both the index test and the reference standard on true positives, false negatives, true negatives and false positives), and certainty of evidence in the summary of findings tables.

For comparative diagnostic accuracy of different index tests, we will evaluate the certainty of the evidence according to the GRADE guidance 31 (45). We will evaluate the same four domains as mentioned above for diagnostic accuracy of individual index tests. We will not consider indirect comparisons (between-study comparisons), given that evidence resulting from indirect comparisons is likely to be of low certainty and the methodological work in this area is under development (45). We will also produce summary of findings tables for comparative diagnostic accuracy if more than two studies are included for each comparison, but will be limited to one outcome only (bilateral tubal occlusion). We will present the number of studies/participants, study design, certainty assessment (risk of bias, indirectness, inconsistency and imprecision), summary of findings (numbers and 95% confidence intervals for both index tests on true positives, false negatives, true negatives, false positives and the differences of these between the two tests), and certainty of evidence in the summary of findings tables.

The supplementary data to this manuscript can be found online at the Cochrane library for study protocols.

REFERENCES

1. Datta J, Palmer MJ, Tanton C, Gibson L.J, Jones KG, Macdowall W, et al. Prevalence of infertility and help seeking among 15 000 women and men.. Human Reproduction September 2016;31(9):2108-2118. [DOI: 10.1093/humrep/dew123]
2. Inhorn MC, Patrizio P. Infertility around the globe: new thinking on gender, reproductive technologies and global movements in the 21st century.. Human Reproduction Update July/August 2015;21(4):411–426. [DOI: https://doi.org/10.1093/humupd/dmv016]
3. Wilkes S, Chinn DJ, Murdoch A, Rubin G. Epidemiology and management of infertility: a population-based study in UK primary care. Family Practice August 2009;26(4):269–274. [DOI: https://doi-org.vu-nl.idm.oclc.org/10.1093/fampra/cmp029]
4. Zhou Z, Zheng D, Wu H, Li R, Xu S, Kang Y, et al. Epidemiology of infertility in China: a population-based study. BJOG March 2018;125(4):432-441. [DOI: 10.1111/1471-0528.14966]
5. Zheng D, Zhou Z, Li R, Wu H, Xu S, Kang Y, et al. Consultation and treatment behaviour of infertile couples in China: a population-based study. Reprod Biomed Online June 2019;38(6):917-925. [DOI: 10.1016/j.rbmo.2018.11.034]
6. Lyons RA, Saridogan E, Djahanbakhch O. The reproductive significance of human Fallopian tube. Human Reproduction Update July/August 2006;12(4):363-372. [DOI: https://doi-org-vu-nl.idm.oclc.org/10.1093/humupd/dml012]
7. Niederberger C, Pellicer A, Cohen J, Gardner DK, Palermo GD, O'Neill CL, et al. Forty years of IVF. Fertility and Sterility July 2018;110(2):185-324. [Other: https://doi.org/10.1016/j.fertnstert.2018.06.005]
8. Committee on Gynecologic Practice, American Society for Reproductive Medicine, Daniel M. Breitkopf, Micah Hill. Infertility Workup for the Women's Health Specialist.. https://www.acog.org/clinical/clinical-guidance/committee-opinion/articles/2019/06/infertility-workup-for-the-womens-health-specialist June 2019;Last access January 10th, 2021.
9. National Institute for Health and Care Excellence. Fertility problems: assessment and treatment. https://www.nice.org.uk/guidance/cg156/chapter/Recommendations#investigation-of-fertility-problems-and-management-strategies February 2013 (update September 2017);Last accessed on January 10th, 2021.
10. Verhoeve HR, Coppus SFPJ, Steeg JW, van der Steures P, Hompes PGA, Bourdrez P, et al. The capacity of hysterosalpingography and laparoscopy to predict natural conception.. Human Reproduction 2011;26(1):134-142. [DOI: 10.1093/humrep/deq263]
11. Tan J, Tannus S, Taskin O, Kan A, Albert AY, Bedaiwy MA. The effect of unilateral tubal block diagnosed by hysterosalpingogram on clinical pregnancy rate in intrauterine insemination cycles: systematic review and meta-analysis. British Journal of Obstetrics and Gynaecology 2018, August;126(2):227-235. [DOI: 10.1111/1471-0528.15457]
12. Ng KYB, Cheong Y. Hydrosalpinx - Salpingostomy, salpingectomy or tubal occlusion. Best practice & research. Clinical obstetrics & gynaecology 2019;59:41-7. [PubMed: 30824209]
13. Melo P, Georgiou EX, Johnson N, van Voorst SF, Strandell A, Mol BWJ, C. et al. Surgical treatment for tubal disease in women due to undergo in vitro fertilisation (Review). Cochrane Database of Systematic Reviews 2020, Issue 10. Art. No.: CD002125. DOI: 10.1002/14651858.CD002125.pub
14. Carson SA, Kallen AN. Diagnosis and Management of Infertility: A Review. JAMA 2021;326(1):65-76. [DOI: 10.1001/jama.2021.4788]
15. Ozkan S, Murk W, Arici A. Endometriosis and Infertility. Epidemiology and evidence based treatments.. Annals of the New York Academy of Sciences 2008;1127:92-100. [DOI: 10.1196/annals.1434.007]

16. Parry JP, Isaacson KB. Hysteroscopy and why macroscopic uterine factors matter for fertility. Fertility and Sterility 2019;112(2):203-211. [DOI: https://doi.org/10.1016/j.fertnstert.2019.06.031]
17. Dreyer K,Out R, Hompes PGA, Mijatovic V. Hysterosalpingo-foam sonography,a less painful procedure for tubal patency testing during fertility workup compared with (serial) hysterosalpingography: a randomized controlled trial. Fertility and Sterility 2014;102(3):821-825. [DOI: http://dx.doi.org/10.1016/j.fertnstert.2014.05.042]
18. Tros R, van Kessel MA, van Kuijk SMJ, Oosterhuis GJE, Kuchenbecker WKH, Kwee J, et al. The capacity of transvaginal hydrolaparoscopy versus hysterosalpingography to diagnose tubal pathology in the work-up of subfertile women, a randomised clinical trial. European Journal of Obstetrics & Gynecology and Reproductive Biology 2019;236:127-132. [DOI: doi.org/10.1016/j.ejogrb.2019.02.035]
19. Wang R, Watson A, Johnson N, Cheung K, Fitzgerald C, Mol BWJ, et al. Tubal flushing for subfertility (Review). Cochrane Database of Systematic Reviews 2020, Issue 10. Art. No.: CD003718. DOI: 10.1002/14651858.CD003718.pub5.
20. 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 Jun;11(6):1204-8. doi: 10.1093/oxfordjournals.humrep.a019356. PMID: 8671424.
21. Maheux-Lacroix S, Boutin A, Moore L, Bergeron M-E, Bujold E, Laberge P, et al. Hysterosalpingosonography for diagnosing tubal occlusion in subfertile women: a systematic review with meta-analysis. Human Reproduction 2014;29(5):953-963. [DOI: 10.1093/humrep/deu024]
22. Ludwin A, Ludwin J, Szczeklik W, Martins WP. Cutaneous small-vessel vasculitis following hysterosalpingo-foam sonography (HyFoSy). Ultrasound in Obstetrics and Gynecology December 2019;54(6):831-834. [DOI: 10.1002/uog.20372]
23. Saunders RD, Shwayder JM and Nakajima ST. Current methods of tubal patency assessment. Fertility and Sterility 2011;95(7):2171-2179. [DOI: 10.1016/j.fertnstert.2011.02.054]
24. Li Y, Qiu J, Ma B, Li T-G, Yi B, Hu H-G, et al. The role of diagnostic magnetic resonance hysterosalpingography in the evaluation of fallopian tubal occlusion of female infertility: A meta-analysis. Clinical imaging April 2021;72:11-18. [DOI: https://doi.org/10.1016/j.clinimag.2020.11.001]
25. Unterweger M, De Geyter C, Fröhlich JM, Bongartz G, Wiesner W. Three-dimensional dynamic MR-hysterosalpingography; a new, low invasive, radiation-free and less painful radiological approach to female infertility. Human Reproduction 01 December 2002;17(12):3138-3141. [DOI: 10.1093/humrep/17.12.3138]
26. Volondat M, Fontas E, Delotte J, Fatfouta I, Chevallier P, Chassang M. Magnetic resonance hysterosalpingography in diagnostic work-up of female infertility – comparison with conventional hysterosalpingography: a randomised study. European radiology February 2019;29:501-508. [DOI: 10.1007/s00330-018-5572-2]
27. Gordts S, Campo R, Rombauts L, Brosens I. Transvaginal hydrolaparoscopy as an outpatient procedure for infertility investigation. Human Reproduction 1998;13(1):99-103. [DOI: 10.1093/humrep/13.1.99]
28. Coenders-Tros R, van Kessel MA, Vernooij MMA, Oosterhuis GJE, Kuchenbecker WKH, Mol BWJ, et al. Performance of outpatient transvaginal hydrolaparoscopy. Human Reproduction 2016;31(10):2285-2291. [DOI: 10.1093/humrep/dew161]
29. Shibahara H, Shimada K, Kikuchi K, Hirano Y, Suzuki T, Takamizawa S, et al. Major complications and outcome of diagnostic and operative transvaginal hydrolaparoscopy. J Obstet Gynaecol Res 2007;33(5):705-709. [DOI: 10.1111/j.1447-0756.2007.00636.x]
30. Federation Medical Specialists. Exploratory Fertility Workup [Oriënterend Fertiliteits Onderzoek]. Richtlijnendatabase. https://richtlijnendatabase.nl/richtlijn/orienterend_fertiliteitsonderzoek_ofo/startpagina_orienterend_fertilitietsonderzoek.html 2015;Last access February 7th, 2021.

31. Coppus SFPJ, Opmeer BC, Logan S, van der Veen F, Bhattacharya S, Mol BWJ. The predictive value of medical history taking and Chlamydia IgG ELISA antibody testing (CAT) in the selection of subfertile women for diagnostic laparoscopy: a clinical prediction model approach. Human Reproduction 2007;22(5):1353-1358. [DOI: 10.1093/humrep/del521]
32. Luttjeboer FY, Verhoeve HR, van Dessel HJ, van der Veen F, Mol BW, Coppus SF. The value of medical history taking as risk indicator for tuboperitoneal pathology: a systematic review. BJOG. 2009 Apr;116(5):612-25. doi: 10.1111/j.1471-0528.2008.02070.x. Epub 2009 Feb 10. PMID: 19220240.
33. Ansari AH. Methylene blue test for assessment of tubal patency: a new and simple technique.. Can Med Assoc J. 1968;99(4):182-184.
34. Shinkins B, Thompson M, Mallett S, Perera R. Diagnostic accuracy studies: how to report and analyse inconclusive test results. BMJ 16 May 2013;346:f2778. [DOI: 10.1136/bmj.f2778]
35. Whiting JF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. Annals of internal medicine 2011;155(8):529-536. [DOI: 10.7326/0003-4819-155-8-201110180-00009]
36. Yang B, Mallett S, Takwoingi Y, Davenport CF, Hyde CJ, Whiting PF, Deeks JJ, Leeflang MMG; QUADAS-C Group†, Bossuyt PMM, Brazzelli MG, Dinnes J, Gurusamy KS, Jones HE, Lange S, Langendam MW, Macaskill P, McInnes MDF, Reitsma JB, Rutjes AWS, Sinclair A, de Vet HCW, Virgili G, Wade R, Westwood ME. QUADAS-C: A Tool for Assessing Risk of Bias in Comparative Diagnostic Accuracy Studies. Ann Intern Med. 2021 Nov;174(11):1592-1599. doi: 10.7326/M21-2234. Epub 2021 Oct 26. PMID: 34698503.
37. Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach.. Journal of Clinical Epidemiology December 2006;59(12):1331-1332. [DOI: 10.1016/j.jclinepi.2006.06.011]
38. Stata [Computer program]. StataCorp. Version 16. College Station, TX: StataCorp LLC: StataCorp, 2019.
39. Freeman SC, Kerby CR, Patel A, Cooper NJ, Quinn T, Sutton AJ. Development of an interactive web-based tool to conduct and interrogate meta-analysis of diagnostic test accuracy studies: MetaDTA. BMC Med Res Methodol 2019 April;19(1):81.. [DOI: 10.1186/s12874-019-0724-x]
40. Review Manager (RevMan) [Computer program]. Version 5.4. The Cochrane Collaboration, 2020.
41. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and Presenting Results.. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. Available from methods.cochrane.org/sdt/hadnbook-dta-reviews. The Cochrane Collaboration, 23 December 2010. [Other: https://methods.cochrane.org/sdt/handbook-dta-reviews]
42. Salameh J-P, Bossuyt PM, McGrath TA, Thombs BD, Hyde CJ, Macaskill P, et al. Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA):explanation, elaboration, and checklist. BMJ 20 May 2020;370:m2632. [DOI: 10.1136/bmj.m2632]
43. Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. Journal of clinical epidemiology 2020;122:129-41. [PubMed: 32060007]
44. Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 2. Test accuracy: inconsistency, imprecision, publication bias, and other domains for rating the certainty of evidence and presenting it in evidence profiles and summary of findings tables. Journal of clinical epidemiology 2020;122:142-52. [PubMed: 32058069]
45. Yang B, Mustafa RA, Bossuyt PM, Brozek J, Hultcrantz M, Leeflang MMG, et al. GRADE Guidance: 31. Assessing the certainty across a body of evidence for comparative test accuracy. Journal of clinical epidemiology 2021;136:146-56. [PubMed: 33864930]

CHAPTER 8

The impact of the COVID-19 pandemic on infertility patients and endometriosis patients in the Netherlands

Kimmy Rosielle‡, Jasmijn Bergwerff‡, Anneke M.F. Scheurs, José Knijnenburg, Bianca de Bie, Jacques W.M. Maas, Annemiek W. Nap, Madelon van Wely, Cornelis B. Lambalk, Mariëtte Goddijn, Inge M. Custers, Laura L. van Loendersloot, Velja Mijatovic

‡These authors contributed equally to this work.

Reprod Biomed Online. 2021 Oct;43(4):747-755.
doi: 10.1016/j.rbmo.2021.06.001. Epub 2021 Jun 12.

ABSTRACT

Research question

How do infertility patients, endometriosis patients and health-care providers rate virtual care as an alternative to physical consultations during the first lockdown of the coronavirus disease 2019 (COVID-19) pandemic in the Netherlands, and how does this influence quality of life and quality of care?

Design

Infertility patients and endometriosis patients from a university hospital and members of national patient organizations, as well as healthcare providers in infertility and endometriosis care, were asked to participate between May and October 2020. The distributed online questionnaires consisted of an appraisal of virtual care and an assessment of fertility-related quality of life (FertiQol) and patient-centeredness of endometriosis care (ENDOCARE).

Results

Questionnaires were returned by 330 infertility patients, 181 endometriosis patients and 101 healthcare providers. Of these, 75.9% of infertility patients, 64.8% of endometriosis patients and 80% of healthcare providers rated telephone consultations as a good alternative to physical consultations during the COVID-19-pandemic. Only 21.3%, 14.8% and 19.2% of the three groups rated telephone consultations as a good replacement for physical consultations in the future. A total of 76.6% and 35.9% of the infertility and endometriosis patients reported increased levels of stress during the pandemic. Infertility patients scored lower on the FertiQol, while the ENDOCARE results care seem comparable to the reference population.

Conclusions

Virtual care seems to be a good alternative for infertility and endometriosis patients in circumstances where physical consultations are not possible. Self-reported stress is especially high in infertility patients during the COVID-19- pandemic. Healthcare providers should aim to improve their patients' ability to cope.

KEY MESSAGE

Patients with infertility, endometriosis patients and their healthcare providers rate telemedicine as a good alternative during the pandemic but agree that it cannot replace physical consultations in the future. Fertility patients report a lower quality of life during this period. Patients with endometriosis judge the care to be comparable to the reference population.

INTRODUCTION

The global outbreak of coronavirus disease 2019 (COVID-19) has led to a significant increase of pressure on healthcare systems all over the world. In the spring of 2020 all elective care and other 'non-essential' medical care was largely restricted or even shut down during the lockdown in the Netherlands in order to prevent the spread of COVID-19 and to focus all resources and healthcare providers on COVID-19 care. For infertility patients and endometriosis patients, this first lockdown period resulted in a temporary cancellation of physical appointments, elective surgery and assisted reproductive technology (ART) during the COVID-19 pandemic in the Netherlands.

In order to maintain continuity of care for both patient groups during the first COVID-19 lockdown, virtual care options such as telephone consultations and video consultations were quickly implemented in most hospitals throughout the Netherlands. Telephone consultations were already being used prior to the pandemic, mainly to communicate the results of diagnostic tests. Video consultations were not widely used in fertility and endometriosis care. With the use of these virtual care alternatives, healthcare providers were able to replace at least a proportion of the cancelled physical appointments in outpatient clinics, thus providing continuity in fertility and endometriosis care.

Under normal circumstances, infertility patients already experience high levels of stress, as well as a high sense of urgency to obtain treatment (1, 2). In addition, patients undergoing fertility treatments show higher levels of depression in comparison to the general population (3, 4). The turbulent period of the first COVID-19 lockdown, with the temporary care restrictions resulting in cancellation of fertility treatments, might have led to additional stress and had a negative impact on the patients' quality of life.

For patients with a chronic disease, such as endometriosis, continuity of care and more specifically the patient-centeredness of the healthcare provided are very important as they are possibly associated with health-related quality of life (5). Patient-centred care is a method of providing care to patients while taking into account 'the preferences, needs and values of the individual patient' (6, 7). The cancellation of physical appointments, elective surgeries and fertility treatments during the COVID-19 lockdown could have a negative impact on the perceived quality of endometriosis care as patients might experience less support from their healthcare providers accompanied by an increase in waiting lists for consultations, surgery and ART.

The aim of this study was to evaluate patient and healthcare provider experiences of the alternative virtual care consultations and to investigate the impact of the restrictive

measures and the shutdown of regular care during the COVID-19 pandemic on fertility-related quality of life and quality of endometriosis care.

MATERIALS AND METHODS

A cross-sectional cohort study was performed in the Netherlands between March 2020 and October 2020. For this study three groups of participants were approached: (1) infertility patients, (2) women with endometriosis and (3) healthcare providers in the field of fertility and/or endometriosis in the Netherlands. In the Netherlands it is very common for gynaecologists to treat both endometriosis and infertility patients. As infertility patients often present with the urgent problem of wishing to conceive and endometriosis patients have complaints and worries of a more chronic nature, both groups can give a unique insight in both current and chronic care while the patients are visiting the same outpatient clinic. The healthcare providers were included in this study to investigate whether patients and professionals shared the same views on virtual care.

Ethical approval

Ethical approval was granted by the institutional review board of Amsterdam UMC for the two respective locations with their own medical ethical review committee (location AMC: reference no. 20.236, approved 7 May 2020; location VUmc: reference no. 2020.264, approved 19 May 2020).

Patient recruitment

To maximize the response, infertility and endometriosis patients were recruited by both Amsterdam UMC, a Dutch university hospital, and by their respective national patient organizations, FREYA (www.freya.nl) and De Endometriose Stichting (www.endometriose.nl). Patients from the university hospital were approached by e-mail when they had an appointment scheduled or were enrolled on a waiting list for ART or elective surgery in Amsterdam UMC between March 2020 and June 2020. Members of both patient organizations were approached via social media, newsletters and blogposts on the websites of the respective patient organizations. Healthcare providers were contacted through the Dutch Society of Obstetrics and Gynaecology (NVOG; www.nvog.nl) as well as the Dutch Society of Fertility Physicians (VVF; www.fertiliteitsartsen.nl). Due to the recruitment via social media, it was not possible to identify unique patients eligible for inclusion. A response rate could therefore not be calculated for the participants from the patients' organizations.

The inclusion criteria for the infertility patients were (i) age ≥18 years, and (ii) women with infertility who were being treated at the Department of Reproductive Medicine of Amsterdam UMC or women who had joined the online network of the national patient organization for infertility. The inclusion criteria for endometriosis patients were (i) age ≥18 years, and (ii) a self-reported endometriosis diagnosis and a member of the national patient organization for endometriosis or receiving treatment at the Endometriosis Centre of Amsterdam UMC. For both groups of patients the exclusion criteria were: (i) age <18 years old, or (ii) an inability to read and write in the Dutch language. Healthcare providers were included if they were a member of the NVOG or the VVF and routinely treated women with infertility and/or endometriosis.

Questionnaires

Three different online questionnaires were developed for infertility patients, endometriosis patients and healthcare providers respectively. The questionnaires were developed in collaboration with the national patient organizations for infertility and endometriosis respectively: FREYA and De Endometriose Stichting. The questionnaires for infertility patients and endometriosis patients were distributed between May 2020 and July 2020. The questionnaire for healthcare providers was distributed between August 2020 and October 2020.

The questionnaires for infertility patients consisted of (i) a demographics and background section, (ii) a section on the assessment of virtual care and stress, and (iii) the Dutch fertility-related quality of life questionnaire (FertiQol). The questionnaires for endometriosis patients consisted of (i) a demographics and background section, (ii) a section on the assessment of virtual care and stress, and (iii) the patient-centeredness of endometriosis care (ENDOCARE) questionnaire (ECQ).

The questionnaire on the assessment of virtual care and stress contained questions on changes in appointments during COVID-19, experience with the different modalities used to alter appointments and care (telephone and video consultations), communication and information during COVID-19, treatment during COVID-19, dealing with change and experiencing stress (Supplementary information). FertiQol is a validated questionnaire evaluating the fertility-related quality of life of infertility patients. It consists of 36 items identifying core quality of life, treatment quality of life and overall quality of life (1, 2). The FertiQoL questionnaire covers six different subdomains: (i) mind–body, (ii) relational, (iii) social, (iv) emotional, (v) environment, and (vi) tolerability (Table 1). Likert scales (0–4) are used to answer the FertiQoL questions, and the outcomes are transformed to a scale ranging from 0 to 100 for all individual subdomains (2). A reference population obtained from Aarts and colleagues was used for a comparison of FertiQoL scores during the COVID-19 pandemic with FertiQoL scores obtained before the pandemic in the Dutch population (1).

Table 1. Examples of questions per FertiQoL domain.

Domain	Example
1. Mind-body	Do you feel drained or worn out because of your fertility problems?
2. Relational	Have fertility problems had a negative impact on your relationship with your partner?
3. Social	Are you socially isolated because of fertility problems?
4. Emotional	Do you feel sad and depressed about your fertility problems?
5. Environment	Are you satisfied with the quality of services available to you to address your emotional needs?
6. Tolerability	Are you bothered by the effect of treatment on your daily or work related activities?

The ECQ is a validated questionnaire evaluating the patient-centeredness of endometriosis care (8, 9). It contains 38 aspects that are assessed using a 4-point Likert scale. Both performance and the importance of the care aspects are rated. The 38 aspects can be divided into 10 categories of patient-centred care: (i) respect for patients' values, preferences and expressed needs; (ii) coordination and integration of care; (iii) information and communication; (iv) physical comfort; (v) emotional support and alleviation of fear and anxiety; (vi) involvement of the significant other; (vii) continuity and transition; (viii) access to care; (ix) technical skills; and (x) endometriosis clinic staff (Table 2). The outcomes are converted to scores ranging from 0 to 100 for each category. The patient-centeredness scores from the same university hospital obtained by Schreurs and colleagues are used as a reference population (10).

The healthcare provider questionnaire consisted of two different subsections: (i) demographics and (ii) assessment of virtual care. The questions used mirrored the questions in the patient questionnaires on virtual care (Supplementary information).

When respondents did not complete the full questionnaire but did complete one or more sections, the completed sections were included in the analysis.

Timeline of COVID-19 restrictions

From 16 March 2020 the Netherlands was in the first lockdown and all elective and non-essential care was paused at that point. Fertility treatments that had started before the 16th of March were completed, but new or subsequent cycles were cancelled. Endometriosis consultations, investigations and surgeries were all cancelled, and only emergency consultations in cases of severe pain or bleeding were possible. From mid-May 2020 planned care was able to slowly restart in the Netherlands.

Table 2. Examples of care aspects per dimension.

Dimension	Example of ECQ Care aspect
1. Respect for patients' values, preferences and expressed needs	My complaints were taken seriously
2. Coordination and integration of care	Care was taken to plan examinations and treatments on 1 day
3. Information and communication	Everything necessary was done so that I would understand the information given
4. Physical comfort	The consultation waiting room is comfortable
5. Emotional support and alleviation of fear and anxiety	I was informed as to the psychological impact of endometriosis
6. Involvement of significant other	There were efforts to involve my partner during consultations
7. Continuity and transition	The physician who is treating me really follows up on my case personally
8. Access to care	I was able to contact a caregiver with specific knowledge of endometriosis in urgent cases
9. Technical skills	I was able to rely on the expertise of the caregivers
10. Endometriosis clinic staff	The caregivers were understanding and concerned during my treatment

The questionnaires for infertility patients and endometriosis patients were sent during this lockdown. The questionnaire for healthcare providers was sent shortly after the lockdown. During this period physical consultations were possible, but only in limited capacity, so telephone and video consultations were still used regularly throughout the Netherlands.

Statistical analysis

Statistical analysis was performed using IBM SPSS Statistics for Windows version 24 (IBM, USA). Descriptive statistics were used to report on the demographics of participants and the assessment of virtual care. One-way analysis of variance were used to test differences between infertility patients, endometriosis patients and endometriosis patients with infertility.

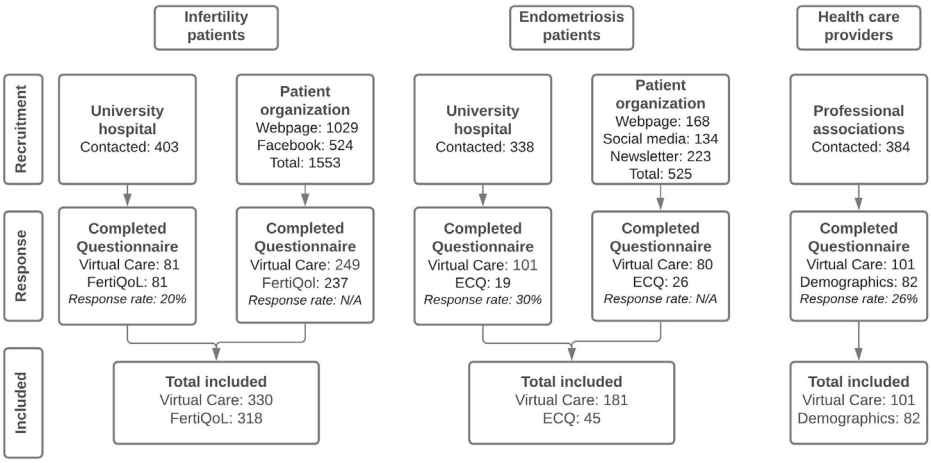
The results of the FertiQoL and ECQ questionnaires were analysed according to their respective guidelines (1, 2, 8, 9). The means and standard deviations provided for the FertiQoL related to the infertility patients were compared with those provided for the Dutch reference population, and mean differences with 95% confidence intervals were calculated. Linear regression was used to assess the association of the baseline variables

age and duration of subfertility with the FertiQoL scores. For the ECQ no comparative statistics were possible in relation to the reference population of the questionnaire, so the results are shown in a bar chart. Answers to the open-ended questions were read and explanations for the results from the questionnaires were sought.

RESULTS

A total of 330 infertility patients (81 from the university hospital, 249 from the patient organization), 181 endometriosis patients (101 from the university hospital, 80 from the patient organization) and 101 healthcare providers responded. Not all questionnaires were fully completed, but all available data were used in the results (Figure 1).

Figure 1. Recruitment of participants.



The response rate was calculated for the participants from the university hospital and the healthcare providers. As the participants from the patient organizations were recruited via social media, the number of individual clicks on the link are given but a response rate cannot be calculated. Virtual care refers to the part of the questionnaire consisting of questions evaluating telephone consultations and video consultations. ECQ, ENDOCARE questionnaire

Patient characteristics

Patient characteristics for the infertility patients are shown in Table 3, and those for the women with endometriosis in Table 4. The median age of the infertility patients was 33. Around half of the infertility patients suffered from primary infertility. The participants from the endometriosis population had a median age of 35 and predominantly reported having moderate to severe endometriosis. One-third of the participants with

endometriosis reported a change in endometriosis complaints during COVID-19. Of these participants, 81.7% reported an increase in endometriosis symptoms.

The healthcare providers had a median age of 45.5 years and the majority were gynaecologists (Supplementary Table 1).

Table 3. Characteristics of infertility participants.

	Value
Age (years), median (IQR]	33.00 (30.00–36.00)
Primary infertility, *n* (%)	167 (50.6)
Has children, *n* (%)	84 (25.5)
Pregnant at time of participation, *n* (%)	1 (0.3)
Duration of infertility (months), median (95% CI)	27.5 (18.0–39.0)

A total of 330 fertility patients completed the patient characteristics part of the questionnaire.

Table 4. Characteristics of endometriosis patients.

	Value
Age (years), median (IQR)	35.00 (31.00–40.50)
Stage of endometriosis[a]	
Minimal to mild	18 (9.9)
Moderate to severe	117 (64.6)
Unknown	46 (25.4)
Surgical confirmation of diagnosis	101 (55.8)
Change in endometriosis-related complaints during COVID-19	60 (33.1)
Reported increase in complaints[b]	49 (81.7)
Reported decrease complaints[b]	16 (26.7)
Hormonal treatment	93 (51.4)
Pregnant at time of participation	3 (1.7)
Has children	61 (33.7)

Data are *n* (%) unless otherwise stated.
A total of 181 endometriosis patients completed the patient characteristics part of the questionnaire.
[a] determined at first diagnosis
[b] Patients were able to report both an increase and a decrease in complaints

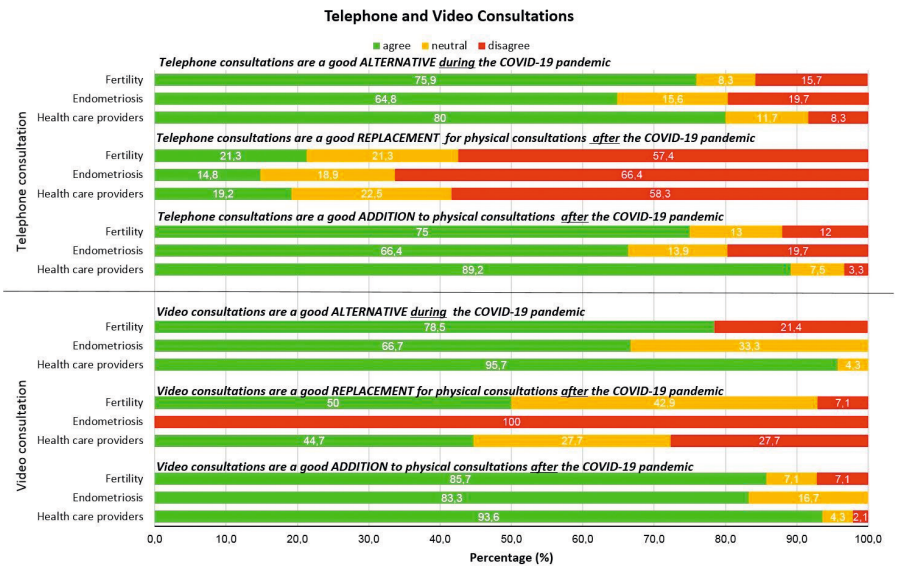
Virtual alternatives to regular care

A total of 88% of infertility patients reported that appointments were cancelled or postponed during the first lockdown of the COVID-19 pandemic. Thirty-three per cent of

participants reported that physical infertility appointments were converted to telephone consultations, and 4% reported conversion to video consultations. Of the endometriosis patients, 67% reported that physical appointments were adjusted to telephone consultations, while 3% of patients had appointments changed to video consultations. Of the healthcare providers, 83% reported that one or more of their physical appointments had been changed to a telephone consultation and 39% reported conversion to video consultations. For both infertility and endometriosis patients, healthcare providers spent a median time of 15 min on telephone consultations and 20 min on video consultations.

The evaluation of virtual care methods by infertility patients, endometriosis patient and healthcare providers is shown in Figure 2. During the lockdown, telephone consultations and video consultations were seen as good alternatives for physical appointments. For the future, both telephone consultations and video consultations were thought to be useful additions to physical appointments. Telephone consultations were not seen as good replacements for future physical appointments by the majority of respondents. On video consultations as a replacement for future physical appointments, respondents were more positive, but still not truly convinced. Endometriosis patients in particular still preferred a physical appointment (six respondents).

Figure 2. Evaluation of virtual care options by infertility patients, endometriosis patients and their healthcare providers.

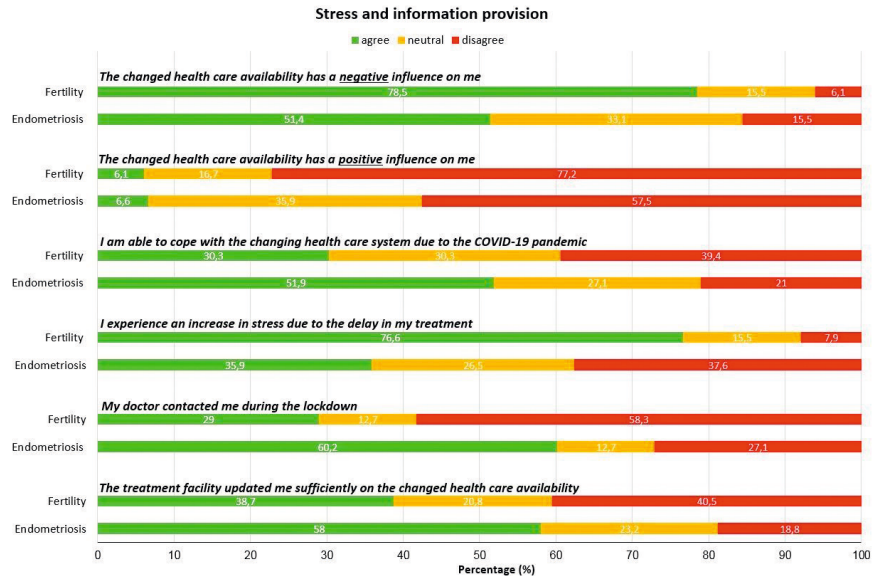


'Good alternative' refers to the situation during the pandemic; 'Good addition' and 'Good replacement' refer to consultations in the time after the COVID-19 pandemic.

Coping with altered care

The results on stress and spread of information as reported by the infertility patients and endometriosis patients are presented in Figure 3. The results on stress ('I experience an increase in stress due to the delay in my treatment') differed between the patient groups: 76.6% of the infertility patients agreed with this statement against only 35.9% of the endometriosis patients ($P < 0.001$). A similar difference was seen in self-reported coping ('I am able to cope with the changing health care system due to the COVID-19 pandemic'), where 30.3% and 51.9% of infertility and endometriosis patients, respectively, agreed ($P < 0.001$). In addition, of a subgroup of endometriosis patients who were currently undergoing fertility treatment ($n = 23$), 60.9% reported increased stress and 43.5% reported that they were able to cope (Supplementary Table 2).

Figure 3. Experienced stress and communication.



A total of 330 fertility patients and 181 endometriosis patients completed the stress and coping-related questions.

Open-ended questions

Both infertility patients and endometriosis patients reported that the use of telephone consultations and video consultations is seen as a feasible option when no physical examinations are needed. For infertility patients, acceptable appointments to use telephone consultations or video consultations for could be communicating laboratory results or solely providing information. Possible examples for the use of telephone consultations and video consultations with endometriosis patients were follow-

up consultations with known patients or discussing alterations in medication. The downside noted by infertility patients during the COVID-19 pandemic is that they missed personal contact with their doctor as fertility treatments are intensive treatments. For endometriosis patients, a reported downside was missing the choice to be able to have a physical consultation when they felt they needed one.

Healthcare providers reported the lack of travel time, being able to provide a safe alternative for healthcare during the pandemic, and time efficiency (e.g. 'patients don't have to wait when the doctor is delayed' and 'more flexible planning of appointments') as benefits of telephone consultations. The additional benefit that video consultations have over telephone consultations according to healthcare respondents is the ability to experience non-verbal communication as well as being able to have conversations with the patient and their partner at the same time. The most important downside of telephone consultations reported by healthcare providers was the lack of non-verbal communication. For video consultations, healthcare providers reported technical difficulties (including connection errors and patients not understanding the technology) to be the most important downside.

Not being able to perform physical examinations and additional investigations (e.g. ultrasonography or blood sampling) and difficulties with providing emotional support were recorded as downsides for both telephone and video consultations.

Fertility patients reported having an increase in stress, reasons being increasing age, which could damage the chance of pregnancy, fear of aggravating underlying illness and ambiguity in information from the hospitals on when treatments could restart.

Infertility patients' quality of life

The fertility-related quality of life information of the infertility patients (n = 318) and the data from a Dutch reference population (n = 473) are shown in Table 5 (1), with the core FertiQoL subdomains shown separately. Although a statistical comparison between the infertility patients in this study and the reference population was not possible due to a lack of access to the data describing the reference population, the quality-of-life scores seem to be lower in the group in the current study compared with the reference population for all subdomains of the FertiQoL.

Table 5. The fertility related quality of life as reported in the core FertiQoL outcome and the subscales.

	Fertility patients mean (SD)	Reference population* mean (SD)	Mean Difference (95% CI)
Core FertiQoL	58.6 (14.8)	70.8 (13.9)	12.2 (10.2 – 14.2)
Social subscale	63.3 (17.8)	74.0 (16.6)	10.7 (8.3 – 13.1)
Relational subscale	71.6 (17.1)	78.2 (14.5)	6.6 (4.4 – 8.8)
Emotional subscale	45.4 (20.2)	59.8 (18.7)	14.4 (11.7 – 17.1)
Mind–body subscale	54.0 (20.1)	70.8 (19.5)	16.8 (13.9–19.6)

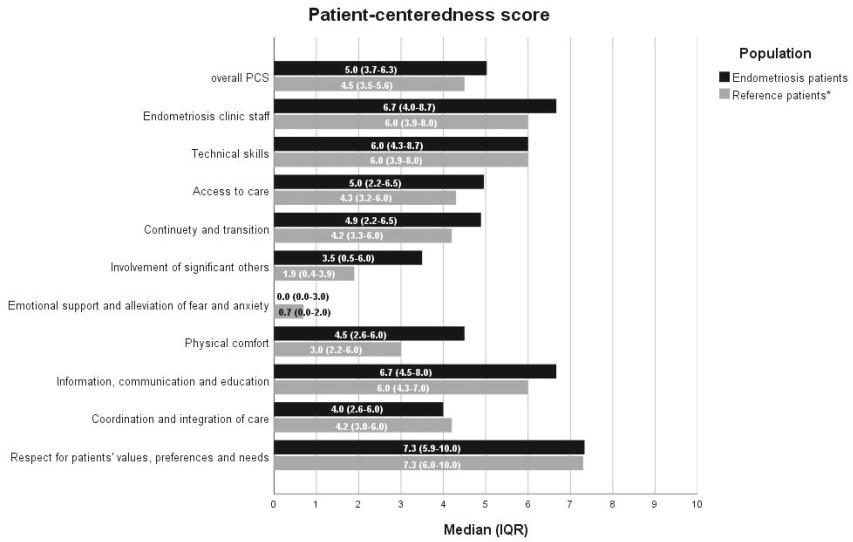
A total of 318 out of 330 patients completed the FertiQoL questionnaire. The reference population consisted of 473 patients.* Aarts et al, 2011

Subgroup analysis showed that increasing female age was associated with a lower relational score ($P = 0.005$) and primary infertility was associated with a higher score on the mind–body and relational domains ($P < 0.001$).

Patient centeredness of endometriosis care

Figure 4 demonstrates the patient-centeredness scores for endometriosis participants ($n = 45$) measured using the ECQ. As a reference, the patient-centeredness scores from 177 patients reported by Schreurs and colleagues (*Schreurs et al., 2020*) were added to the figure as a comparison with the pre-COVID-19 situation. The patient-centeredness of endometriosis care during the COVID-19 pandemic seems comparable to that of the reference population that was used.

Figure 4. Patient-centeredness scores (PCS) by dimension, as measured by the ENDOCARE questionnaire.



A total of 45 out of 181 endometriosis patients completed the ENDOCARE questionnaire. The reference population consisted of 177 patients.
* Schreurs et al, 2020. IQR, interquartile range.

DISCUSSION

This study shows that the use of virtual care, specifically telephone and video consultations, during the lockdown caused by the COVID-19 pandemic proved to be a good alternative to regular physical consultations for the large majority of patients with infertility and endometriosis and their healthcare providers. Both the patient groups and the healthcare providers thought that the use of telephone consultations would be a good addition to regular care in the future, but that it could not replace regular physical consultations. All groups were positive about video consultations, although video consultations had not yet been widely implemented at the time of this study. Quality of life in infertility patients appeared to be lower for all subdomains when compared with the reference population. The patient-centeredness of endometriosis care during the COVID-19 pandemic seems comparable to that of the reference population used.

The first lockdown in the Netherlands came quite suddenly. One of the strengths of this study was the early distribution of questionnaires to the patients during the first lockdown of the COVID-19 pandemic, which limits the chance of recall bias from patients. The use of the validated FertiQoL and ECQ allowed for an objective and validated measurement of quality of life for infertility patients and of patient-centeredness of endometriosis care during COVID-19.

The questionnaires were developed in collaboration with two patient organizations to ensure that the questions were relevant and reflected patients' experiences during that stage of the pandemic. Due to the short time frame of this study and despite the extensive collaboration with the patient organizations and multiple reminders to complete the questionnaires, the response rate remained relatively low, and this is a potential source of response bias.

The use of virtual care as an alternative for physical consultations during the pandemic was rated positively by patients; these results are in line with recently published studies during the COVID-19 pandemic (11-14). Yet the replacement of physical consultations by telephone consultations in the future was not seen as a desirable option by the majority of patients from both groups. A possible explanation is that fertility treatments are not possible without physical appointments, for instance for the monitoring of follicle growth. Endometriosis patients receive regular check-ups where their physician routinely performs a physical examination, including transvaginal ultrasonography. The desire to obtain reassurance in this way might also explain why endometriosis patients prefer physical appointments. However, the results should be interpreted with caution, as the number of respondents to these questions was low.

In the current study, the use of video consultations was limited in both patient groups while 39% of the healthcare providers reported using video consultations. A possible reason for this difference could be that the questionnaire for healthcare providers was distributed 3 months later than the questionnaire for patients. After the first lockdown an increased use of video consultations may have occurred as hospitals were developing strategies to continue consultations without inviting patients to their clinics. Another possibility for this difference is that healthcare providers have multiple appointments a day, so an overestimation of the number of video consultations by recall bias cannot be excluded. In accordance with both patient groups, the healthcare providers reported that video consultations are a good addition to regular care for the future, and this is also in line with other recent studies (15, 16). The benefit of video consultations compared with telephone consultations lies in the visual aspect, which aids non-verbal communication and gives a more personal interaction (16).

During the same period that this study was being conducted, the Dutch Institute for Public Health and the Environment reported that 23.9% of Dutch citizens were experiencing high levels of stress (17). This is much lower than the 76.7% of the infertility patients who reported stress in the current study. Unfortunately, the specific reasons for this increase of stress were not explored. Earlier studies showed that women with infertility experience a high sense of urgency to obtain treatment (1). The current delay in treatment due to the pandemic could intensify feelings of stress and urgency, as treatment cancellation has previously been negatively associated with quality of life in infertility patients (18). A recent study by Boivin and colleagues found similar results: 11% of participants reported feeling unable to cope with the stress caused by fertility clinic closure (19). Another study investigating the perceptions and psychological impact of the COVID-19 pandemic on infertile patients identified that feeling helpless and having lower self-control and less social support were correlated with higher psychological distress (20).

In the current study, women with endometriosis experienced less stress than those with infertility. This may be related to the chronic nature of their illness in comparison to the more time-sensitive issues that patients with fertility problems face. In addition, even though continuous endometriosis care is valued as important, endometriosis patients might be able to accept a temporary decrease of care possibilities or a delay in their yearly appointment. A study performed in Turkey during the first peak of the COVID-19 pandemic showed that 83.9% of responders were afraid of experiencing endometriosis-related problems during the pandemic, and 63.0% were afraid that their healthcare professional might be unavailable to them during the pandemic (21). In the current study, only 33.1% of patients actually experienced changes in endometriosis-related complaints,

indicating that the high levels of fear of endometriosis-related problems as previously reported by patients are an overestimation of the actual numbers.

The results of the patient-centeredness scores reported by women with endometriosis during the first lockdown were similar to those of the reference population outside the COVID-19 pandemic. This indicates that even during a COVID-19 pandemic, the same care aspects remain important to patients.

At the time of writing, the Netherlands is recovering from the second lockdown of the pandemic. In contrast to the first lockdown, fertility treatments and endometriosis care have continued, with some restrictions on the number of physical consultations a day and a diminished capacity for surgical and ART care. With the results of the current study in mind, the importance of continuity of care can be underlined. Even though fertility care can be classified as 'non-essential' or 'not life threatening' during the COVID-19 pandemic, this study shows that restriction of care is associated with an increase in stress and a lowered quality of life among infertile women. It is necessary to stress the importance of the use of virtual care in combination with regular physical care to continue treatment for infertility patients as much as is possible.

This study shows that women with endometriosis do not experience the same level of stress as a result of the temporary halting of their treatment as women with infertility do, and the ECQ results are comparable to the reference population. It can, however, be advised that healthcare providers should be accessible for endometriosis patients, and that they should make sure that their patients know how to reach them with questions related to an increase of endometriosis complaints.

CONCLUSIONS

Virtual care seems to be a good alternative for infertility and endometriosis patients in circumstances where physical consultations are not possible. Self-reported stress is especially high in infertility patients during the COVID-19-pandemic and they do not feel that they can cope well with the changes to their care. Healthcare providers should aim to increase their patients' ability to cope with the healthcare changes. Future research should focus more on the role of video consultations as this approach has only recently been implemented in current care.

REFERENCES

1. Aarts JW, van Empel IW, Boivin J, Nelen WL, Kremer JA, Verhaak CM. Relationship between quality of life and distress in infertility: a validation study of the Dutch FertiQoL. Hum Reprod. 2011;26(5):1112-8.
2. Boivin J, Takefman J, Braverman A. The fertility quality of life (FertiQoL) tool: development and general psychometric properties. Hum Reprod. 2011;26(8):2084-91.
3. Massarotti C, Gentile G, Ferreccio C, Scaruffi P, Remorgida V, Anserini P. Impact of infertility and infertility treatments on quality of life and levels of anxiety and depression in women undergoing in vitro fertilization. Gynecol Endocrinol. 2019;35(6):485-9.
4. Volgsten H, Skoog Svanberg A, Ekselius L, Lundkvist O, Sundstrom Poromaa I. Prevalence of psychiatric disorders in infertile women and men undergoing in vitro fertilization treatment. Hum Reprod. 2008;23(9):2056-63.
5. Apers S, Dancet EAF, Aarts JWM, Kluivers KB, D'Hooghe TM, Nelen W. The association between experiences with patient-centred care and health-related quality of life in women with endometriosis. Reprod Biomed Online. 2018;36(2):197-205.
6. Geukens EI, Apers S, Meuleman C, D'Hooghe TM, Dancet EAF. Patient-centeredness and endometriosis: Definition, measurement, and current status. Best Pract Res Clin Obstet Gynaecol. 2018;50:11-7.
7. WHO. Quality of care: a process for making strategic choices in health systems. In: Organization WH, editor. 2006.
8. Dancet EA, Ameye L, Sermeus W, Welkenhuysen M, Nelen WL, Tully L, et al. The ENDOCARE questionnaire (ECQ): a valid and reliable instrument to measure the patient-centeredness of endometriosis care in Europe. Hum Reprod. 2011;26(11):2988-99.
9. Dancet EA, Apers S, Kluivers KB, Kremer JA, Sermeus W, Devriendt C, et al. The ENDOCARE questionnaire guides European endometriosis clinics to improve the patient-centeredness of their care. Hum Reprod. 2012;27(11):3168-78.
10. Schreurs AMF, van Hoefen Wijsard M, Dancet EAF, Apers S, Kuchenbecker WKH, van de Ven PM, et al. Towards more patient-centred endometriosis care: a cross-sectional survey using the ENDOCARE questionnaire. Hum Reprod Open. 2020;2020(3):hoaa029.
11. Barsom EZ, Feenstra TM, Bemelman WA, Bonjer JH, Schijven MP. Coping with COVID-19: scaling up virtual care to standard practice. Nat Med. 2020;26(5):632-4.
12. Chaudhry H, Nadeem S, Mundi R. How Satisfied Are Patients and Surgeons with Telemedicine in Orthopaedic Care During the COVID-19 Pandemic? A Systematic Review and Meta-analysis. Clin Orthop Relat Res. 2020.
13. Li HL, Chan YC, Huang JX, Cheng SW. Pilot Study Using Telemedicine Video Consultation for Vascular Patients' Care During the COVID-19 Period. Ann Vasc Surg. 2020;68:76-82.
14. Lun R, Walker G, Daham Z, Ramsay T, Portela de Oliveira E, Kassab M, et al. Transition to virtual appointments for interventional neuroradiology due to the COVID-19 pandemic: a survey of satisfaction. J Neurointerv Surg. 2020;12(12):1153-6.
15. Barsom EZ, Jansen M, Tanis PJ, van de Ven AWH, Blusse van Oud-Alblas M, Buskens CJ, et al. Video consultation during follow up care: effect on quality of care and patient- and provider attitude in patients with colorectal cancer. Surg Endosc. 2020.
16. Jimenez-Rodriguez D, Santillan Garcia A, Montoro Robles J, Rodriguez Salvador MDM, Munoz Ronda FJ, Arrogante O. Increase in Video Consultations During the COVID-19 Pandemic: Healthcare Professionals' Perceptions about Their Implementation and Adequate Management. Int J Environ Res Public Health. 2020;17(14).

17. RIVM. Resultaten 3e ronde: Welbevinden en leefstijl 2020 [Available from: https://www.rivm.nl/gedragsonderzoek/maatregelen-welbevinden/resultaten-3e-ronde-gedragsonderzoek/resultaten-3e-ronde-welbevinden-en-leefstijl.
18. Heredia M, Tenias JM, Rocio R, Amparo F, Calleja MA, Valenzuela JC. Quality of life and predictive factors in patients undergoing assisted reproduction techniques. Eur J Obstet Gynecol Reprod Biol. 2013;167(2):176-80.
19. Boivin J, Harrison C, Mathur R, Burns G, Pericleous-Smith A, Gameiro S. Patient experiences of fertility clinic closure during the COVID-19 pandemic: appraisals, coping and emotions. Hum Reprod. 2020;35(11):2556-66.
20. Ben-Kimhy R, Youngster M, Medina-Artom TR, Avraham S, Gat I, Haham LM, et al. Fertility patients under COVID-19: Attitudes, Perceptions, and Psychological Reactions. Hum Reprod. 2020.
21. Yalcin Bahat P, Kaya C, Selcuki NFT, Polat I, Usta T, Oral E. The COVID-19 pandemic and patients with endometriosis: A survey-based study conducted in Turkey. Int J Gynaecol Obstet. 2020;151(2):249-52.

SUPPLEMENTARY MATERIAL

Supplementary Table 1. Characteristics of health care providers

	Healthcare providers n= 82
Age median (median[IQR])	45.5 [39.0-56.0]
Profession n(%)	
Gynaecologist	46 (56.1)
Fertility physician	35 (43.9)
Expertise n(%)	
Reproductive health care	49 (59.7)
Endometriosis	8 (9.8)
Both	25 (30.5)

Supplementary Table 2. Levels of stress, coping, positive effect and negative effect for infertility patients, endometriosis patients and endometriosis patient receiving ART

	Infertility (n=330)	**Endometriosis (n=158)**	**Endometriosis in ART (n=23)**	**p-value**
Increase in stress n(%)				p=0.000
Agree	253 (76.7)	51 (32.3)	14 (60.9)	
Neutral or disagree	77 (23.3)	107 (67.7)	9 (39.1)	
Coping n(%)				p=0.000
Agree	100 (30.3)	84 (53.2)	10 (43.5)	
Neutral or disagree	230 (69.7)	74 (46.8)	13 (56.5)	
Positive effect n(%)				p=0.861
Agree	20 (6.1)	11 (7.0)	1 (4.3)	
Neutral or disagree	310 (93.9)	147 (93)	22 (95.7)	
Negative effect n(%)				p=0.000
Agree	259 (78.5)	80 (50.6)	13 (56.5)	
Neutral or disagree	71 (21.5)	78 (49.4)	10 (43.5)	

The questionnaires can be found as supplementary information on RBM Online.

CHAPTER 9

General discussion

GENERAL DISCUSSION

In this final chapter we will use the results of the studies presented in previous chapters to answer the research questions that were raised in the general introduction of this thesis. We will place these results in the current field of science and outline implications for clinical practice as well as the perspectives for future research.

Summary

A. HSG: a diagnostic procedure with therapeutic effect

Although HSG was introduced as a diagnostic test, the fertility enhancing effect of HSG with oil-based contrast has been debated for decades. This therapeutic effect of an HSG with oil-based contrast in couples with unexplained infertility has been established by the clear results of the H2Oil study (2) and its recent replication (1) as well as subsequent systematic reviews and meta-analyses (3, 4). Further research questions regarding the mechanism of action, the efficacy in other groups of infertile couples and the safety were raised.

I. Does the fertility enhancing effect of tubal flushing during HSG change over time?

We analysed the follow-up data from the H2Oil study up until three years after HSG and calculated consecutive hazard ratios for the chance of the occurrence of an ongoing pregnancy for the group that had received oil-based contrast versus the group that had received water-based contrast. Our results in **Chapter 2** show that the hazard ratio for ongoing pregnancy was 1.71 with of oil-based contrast immediately after HSG and reduced to 1 at approximately 2 years after tubal flushing. In other words, the treatment effect of oil-based over water-based contrast during HSG lasts approximately two years.

II. What is the effectiveness and cost-effectiveness of the use of oil-based versus water-based contrast medium during HSG in terms of live birth in women who are 39 years or older, women who have a high risk for tubal pathology or who have an ovulation disorder?

We have set up the study 'H2Oil2' to answer this question in **Chapter 3**. The study commenced in 2019 and is currently including women in multiple academic and non-academic hospitals throughout the Netherlands and the United Kingdom. With the results of this study, we hope that a new (inter)national guideline on the use of HSG

during fertility work-up can be ascertained including evidence-based advice tailored to the largest subgroups of infertile couples.

We strive to provide tailored advice when it comes to fertility treatments as well. The treatment advice can range from expectant management, meaning trying to conceive naturally, up to IVF/ICSI treatment. While for couples with severe male infertility ICSI is often the only option, it can be unclear what the best treatment is for couples with unexplained infertility and especially when to start treatment. The first-line treatment for couples with unexplained infertility is often intra uterine insemination with mild ovarian stimulation (IUI-MOH) (5). It is unclear when to start treatment after expectant management. Previous research showed that the effectiveness of IUI-MOH versus expectant management varies based on the prognosis for natural conception as calculated by the Hunault model (6).

It is important to replicate such findings before implementing them into clinical practice. Therefore, we formulated the following research question:

III. *Can we replicate the finding that the benefit of IUI with mild ovarian stimulation compared to expectant management for couples with unexplained infertility depends on the prognosis of natural conception?*

Using the H2Oil database, we were able to compare pregnancy chances for couples in expectant management and couples undergoing IUI-MOH treatment in **Chapter 4**. We did not find a strong correlation between the prognosis of natural conception and the effect of IUI-MOH on pregnancy chances. We did show that IUI increases the chance of conception when compared to expectant management. This treatment effect was dependent on the timing of starting IUI-MOH. In couples with a longer duration of infertility, the benefit of IUI-MOH over expectant management was greater. This leads to our advice to postpone the start of IUI-MOH after the diagnosis of unexplained infertility with several months.

With the fertility-enhancing effect of tubal flushing with oil-based contrast, HSG has moved from a purely diagnostic procedure to a therapeutic intervention. Most probably, once the therapeutic effect is included in fertility guidelines, the HSG with oil-based contrast will be performed more often in women with unexplained infertility. For that reason, it is even more important to consider the adverse effects of the procedure.

To provide a complete overview of the adverse effects we wanted to answer the following question:

IV. What is the frequency and what are the consequences of complications during and after an HSG with oil-based contrast used in infertile patients?

The review we performed to answer this question was presented in **Chapter 5**. After thorough examination of all published literature since the 1920's regarding the use of HSG for infertility, the most reported complication was found to be intravasation of contrast, with a pooled overall frequency of 2.7% (95%CI 1.7 – 3.8). This complication led to serious consequences in four cases , out of the 19,339 HSG procedures that were reported in literature.

The second most reported complication is infection, occurring in 0.9% of all reported HSG's with oil-based contrast (95% CI 0.47 – 1.50) and 1.9% after HSG with water-based contrast (95% CI 0.27–4.60). Another complication of interest relates to the iodine-content of the used HSG contrast. Women with previous subclinical hypothyroidism are more prone to developing overt hypothyroidism after HSG, with a prevalence of 35.7% versus 0–2.2% in euthyroid women (7). Two studies investigated thyroid function in offspring born after an HSG (8, 9). The first, a Japanese studie, showed abnormal neonatal thyroid screening in 2.4% of the children born after HSG. A retrospective analysis of the Dutch children born to mothers that had participated in the H2Oil study showed no abnormalities at the heel prick screening test in the first week after birth.

Seven reports were identified that described fatal sequelae of an HSG in infertility, all between 1928 and 1950. The presumed causes of death were infection (prior to standard availability of antibiotics) and allergy. We concluded that safety concerns should not be a reason to withhold tubal flushing with oil-based contrast during HSG in women facing infertility.

B. Improving HSG experience

The HSG is known to be an uncomfortable and/or painful procedure, and analgesia that is both effective and appropriate is lacking (10).

Virtual Reality (VR) is emerging as an alternative to traditional pharmacological analgesics. It diverts attention away from physical sensations by bringing the user into a virtual world (11). Its effectiveness has been proven in medical procedures in various fields (12). This led to the following research question:

Can Virtual Reality reduce pain or discomfort during and after HSG?

In our randomized controlled trial, presented in **Chapter 6**, we demonstrated that VR is not effective at reducing peak pain or overall pain during HSG when compared to a control group (VAS peak pain 7.0 vs 6.6, p=0.40 and overall pain 5.0 vs 4.9, p=0.80). Time spend thinking about pain during the procedure was also rated similar in the two groups (3.26 vs 3.52 on scale 1-5, p=1.72).

Prior to the procedure, most patients (77%) expected that VR would be able to distract them from the procedure. Only 19% of patients expected that VR would be able to reduce their discomfort.

Women in the intervention group scored their level of distraction as mean 3.54 on a scale of 1-5, but scored their level of being immersed into the VR lower at mean 2.85. Once women were asked what type of VR surrounding they would prefer, most women mentioned that they thought a comedy series or more engaging nature films with baby animals would work best as a distraction.

C. Alternatives to HSG

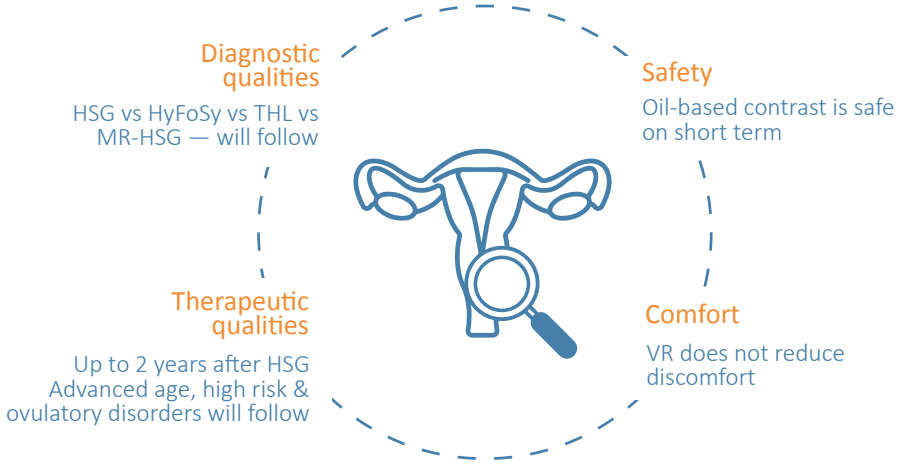
The discomfort of HSG is one of the reasons for clinicians to look for alternative tubal patency tests. Other disadvantages of HSG include the use of ionizing radiation and iodine containing contrast, the need for a radiology department and a technician, and the costs. Alternative visual tubal patency tests have been developed over the last decades that are more patient friendly, do not require X-ray techniques and are less costly. Nowadays, different modalities are used for the evaluation of tubal patency: ultrasound (hysterosalpingo-foam-sonografy (HyFoSy), hysterosalpingo-contrast-sonography with use of saline (HyCoSy), both in 2D and 3D imaging), MRI (MR-HSG) and transvaginal hydrolaparoscopy (THL). All these alternatives to HSG have their own advantages and disadvantages, and their implementation in fertility work-up varies greatly between countries and regions. Studies have been performed to compare the tests to each other or to HSG, in order to establish the correct position for each of the tests within the fertility work-up.

However, the diagnostic accuracy of all these types of tubal patency tests has not been compared to each other and to the current reference standard, laparoscopy and dye testing, systematically. We are currently undertaking a Cochrane review, of which the protocol can be found in **Chapter 7**, to answer the question:

> *What is the diagnostic accuracy of the various types of visual tubal patency tests for diagnosing tubal occlusion?*

The important aspects of tubal patency testing and the answers that this thesis provides are summerized in Figure 1.

Figure 1. Aspects of tubal patency testing



Zooming in on the figure from the introduction, visualizing the route that infertile couples take in their wish to conceive, we highlight the most important aspects of tubal patency testing. With the results of the articles presented in this thesis we can conclude that the therapeutic effect of oil-based contrast is present up to two years after HSG, that oil-based contrast during HSG is safe, and that VR is not effective at reducing discomfort during HSG. We have yet to investigate the diagnostic qualities of HSG in comparison to other visual tubal patency tests, and the effect of oil-based contrast during HSG in women with other types of infertility than idiopathic subfertility or subfertility based on mild male factor.

D. COVID-19 and infertility

Coronavirus Disease 2019 (COVID-19) reached the Netherlands during the course of this project at the end of February 2020 and gave rise to the shutdown of all non-essential health care in March 2020. This led to the temporary closure of all outpatient clinics for infertility and endometriosis in the Netherlands for a period of 3 months. During this time, many clinics tried to set up some form of eHealth for their patients to be able to help at least some of them. This gave rise to the final research question addressed in this thesis:

> *What is the impact of the treatment pause on quality of life and quality of care of patients with infertility or endometriosis in the Netherlands, and how do they rate eHealth?*

In our prospective questionnaire-based study in **Chapter 8** we observed that patients facing infertility scored lower in terms of quality of life when compared to a Dutch

reference population of infertile women (13). The patients from our study additionally reported an increase in stress due to the delay in treatment. Only a third of the infertile patients (30.3%) declared to be able to cope with the change in availability of healthcare during the pandemic, whereas over half (51.9%) of the patients with endometriosis were able to cope. The endometriosis patients showed less detrimental effects and signs of the COVID-19 pandemic to their endometriosis-related health. 'Only' 35.9% of the endometriosis patients reported to experience an increase in emotional stress due to treatment delay, compared to 76.6% of the infertility patients that experienced an increase in stress. Of the 33.1% of endometriosis patients that noticed a change in their endometriosis related complaints, 81.7% reported an increase of symptoms. In terms of quality of care for endometriosis patients, the same aspects were imported to them as in a pre-pandemic cohort (14).

Patients reported that they preferred the telephone appointments that were used during the lockdown as an addition to regular care, but they did not consider them as a good replacement of physical visits to the outpatient clinic. Figure 2 summarizes the impact of COVID-19.

Figure 2. Effects of COVID-19

IMPACT OF COVID-19

Infertility:	**Endometriosis:**
↓ Quality of life	= Quality of life
↓ Psychological stress	= Psychological stress

The most important outcomes of the study on the effects of COVID-19 for infertility and endometriosis patients are inserted in the image from chapter 1.

Comparison to current literature & future perspectives

A. HSG: a diagnostic procedure with therapeutic effect

A replication of the H2Oil study (2) was recently performed in China (1). At 6 months follow-up the ongoing pregnancy rate was 29.1% in the group that had received oil-based contrast versus 20.1% in the water-based contrast group (RR 1.44, 95% CI 1.15 to 1.81, p=0.001), confirming the results previously found regarding ongoing pregnancy rates in the H2Oil study (2).

Despite the similar inclusion criteria and comparable design, there are also some notable differences between this replication study and the Dutch H2Oil study. In the study by Zhang et al, the ongoing pregnancy rates were much lower (29.1% after oil-based contrast and 20.1% after water-based contrast) than in the H2Oil study (39.7% after oil-based contrast and 29.1% after water-based contrast), and the pain scores were also much lower. The differences in ongoing pregnancy rates between the two studies might be explained by the proportion of primary infertility (67.4% in H2Oil study and 83.3% in the replication study). In any case, this study provides critical information for health care providers all over the world; due to differences in availability, different types of oil-based and water-based contrast were used in this study compared to the H2Oil study. This means that the treatment effect is most likely independent of the type of oil-based or water-based used.

The Cochrane review from 2020 already showed clear results on the fertility enhancing effect of oil-based versus water-based contrast (4). Addition of the replication study to a systematic review on this topic will most likely create more certainty of evidence and a smaller confidence interval, making the results even more robust.

The long-term follow-up of the H2Oil study demonstrated that the fertility enhancing effect of oil-based contrast is long lasting and the effect is still visible over a time span of 5 years, with a significant difference in live births in favour of the use of oil contrast (RR 1.11; 95% CI 1.03-1.20) (15). The cost effectiveness analysis showed that the overall costs between oil-based contrast and water-based contrast are comparable and therefore an HSG with oil-based contrast is deemed the better choice in the evaluated population of couples with unexplained infertility or subfertility based on a mild male factor (16).

It is unclear whether the fertility enhancing effect and its cost-effectiveness holds for other populations and especially for couples with other types of infertility. It is possible that within the group of couples with unexplained infertility, there is a large group of women with a yet unidentified pregnancy hindering factor that is remedied by use of oil-based contrast during HSG. All other couples have at least one identified factor that reduces their chance of conception and it is thinkable that these couples might benefit less from the oil-based contrast as the role of the unidentified pregnancy hindering factor might be smaller.

While the fertility enhancing effect of oil-based contrast has been proven in women with unexplained infertility, this is not implemented into practice in all centres in the Netherlands. The Dutch guideline for infertility work-up dates from 2015 and describes the HSG as a diagnostic test, without taking into account the fertility enhancing effect

of the different types of contrast (17). It restricts the performance of tubal testing by HSG to women with a high risk for tubal pathology based on their medical history. While some hospitals adhere strictly to this guideline, others have changed their local policy based on the results of the H2Oil study and the results of the subsequent systematic reviews and meta-analyses. Therefore, numerous local policies can be encountered in the Netherlands; some hospitals perform an HSG as a standard test in their fertility work-up and always use either water-based or oil-based contrast, other sites choose their type of contrast based on the results of the H2Oil study or based on the perceived risk for tubal pathology. And some hospitals rarely perform HSG at all. This means that the HSG practice varies greatly throughout the Netherlands which is undesirable. We hope that the results of the studies in this thesis will aid in the formation of new national and international guidelines.

MECHANISM

Our results on the duration of the fertility enhancing effect of tubal flushing with oil based contrast support the hypothesis of mechanical flushing. In this hypothesis it is thought that oil-based contrast has a greater potential of flushing out pregnancy-hindering debris from the Fallopian tubes, an effect that would lead to a (semi-) permanent increase in fecundity. The duration of the fertility enhancing effect was even longer in the subgroup of patients reporting a higher pain score during HSG, further strengthening this hypothesis. In contrast, the hypothetical mechanism in which the endometrial lining is affected by the use of oil-based contrast would most likely be more temporary as the endometrium renews with every menstrual cycle. Future research should focus on the pressure of infusion of different types of contrast and whether or not a drop of pressure is detectible. This could help to determine whether mechanical flushing is indeed the most likely mechanism of action and whether lubrication of the tubal cilia is involved as well.

SAFETY

Intravasation

The most feared complication of HSG, intravasation, occurs only in a very small percentage of cases based on our literature review (2.7% of HSGs with use of oil-based contrast). In the H2Oil study, intravasation was only reported once in all 1119 HSGs. Radiological techniques have changed drastically over the past five decades, reducing the dosage of iodizing radiation while creating images of higher quality. Fluoroscopy, enabling the physicians and technicians to view live imaging during the procedure, has enabled early diagnosis of intravasation making sure the procedure can be halted before serious consequences appear. It should be noted that in most studies there was no

predefined definition of intravasation or oil emboli and therefore 44 of the 48 studies had to be classified as 'high risk of bias' regarding the measurement of the occurrence of intravasation.

Thyroid (dys) function

We've noted that the maternal and foetal thyroid function after HSG is receiving more attention in the past years, both in literature and in clinical practice. The iodine content in oil-based contrast is 480IU/l, while the iodine content of the various types of water-based contrast that are used for HSG ranges from 270-320IU/l. Additionally, oil-based contrast resorbs at a much slower rate than water-based contrast meaning that the exposure to excess iodine is much longer when oil-based contrast is used (18). Iodine affects the thyroid gland and both cases of hyperthyroidism and hypothyroidism following the use of iodinated contrast have been described, the latter being more common. Thyroid abnormalities post-HSG are most prevalent in women displaying subclinical hyperthyroidism or subclinical hypothyroidism prior to HSG (7). A recent prospective cohort study showed that 38% of the 188 previously euthyroid women developed subclinical hypothyroidism in the weeks after an HSG with oil-based contrast, and nearly all (98%) women showed iodine excess as defined by urinary iodine contents (19). The mean dosage of oil-based contrast was 6.2ml (standard deviation 2.6), a relation between the dosage and thyroid function post-HSG was unfortunately not described in the paper. An older study describes the occurrence of hypothyroidism in 22.5% of the 31 women one month after an HSG with oil-based contrast, versus 9.5% of the 42 women one month after an HSG with water-based contrast (20). In the study by Mathews et al the TSH levels dropped back to normal around 12-16 weeks on average. These studies underline the importance of a comparison in thyroid abnormalities after an HSG with oil-based contrast versus water-based contrast. Additionally, as the authors conclude, the effect of temporary thyroid dysfunction on pregnancy chances and foetal development should be analysed further.

A foetus is dependent on its mothers thyroid hormones for development during the first trimester, the foetal thyroid gland starts endogenous thyroid hormone production around week 16 (21). Intra-uterine deprivation of thyroid hormone impairs foetal neurological development (22). Persistent maternal iodine excess could potentially lead to congenital hypo- or hyperthyroidism in the neonate, although retrospective analysis of infants born to mothers participating in the H2Oil study showed no thyroid abnormalities in the heel-prick screening of the neonates (9, 23). Despite these reassuring results, it is still possible that these infants experienced temporary hypo- or hyperthyroidism in utero with unknown sequelae.

Oil-based contrast can also be used in neonates and young children for cardiac procedures, in which case the contrast is injected intravenously. The FDA has recently issued a warning regarding the use of iodine-containing contrast intravenously in children as studies have linked this to a decreased functioning of the thyroid gland in 1-15% (24). The use of iodine-containing contrast during pregnancy is not mentioned in this warning. Long term follow-up of children born after an HSG with oil-based contrast is warranted to rule out an effect of (temporary) maternal thyroid dysfunction in the first trimester on neuro-cognitive development later in their offspring. Children born to mothers who participated in the original H2Oil study (3) are currently participating in a long-term follow-up study, the Neuro H2Oil study, in which neuro-cognitive development is investigated. The results are expected very soon.

B. Improving HSG experience

Virtual Reality has been proven an effective method for pain relief during various medical procedures (12). However, there are also several published studies reporting no significant effect of VR (25, 26) in line with the results of our study. Several points have been raised that can explain these contradicting findings. One of the explanations is the level of immersion or presence into the Virtual Reality. A higher level of immersion or presence indicates a greater feeling of actually being inside the VR, and this poses a greater distraction from physical sensations (27). The mean score of immersiveness ('Experiencing to be inside the VR environment') in our intervention group was 2.9 on a scale of 1.0-5.0, indicating that this potentially was suboptimal. This relatively low level of immersion can be caused by the lack of sound and subsequently being able to hear the physician and radiology technician. Additionally, the duration of the HSG procedure is longer than the duration of the movies available on the VR glasses. Once a movie ends, the user has to choose or wait for a new movie to start and this brings the user back into the 'real world'. A study in which patients underwent a short (mean <4 minutes) diagnostic hysteroscopy showed a positive effect of VR (28), while another study in which the hysteroscopy took longer (median around 8 minutes) showed no effect (26).

Yet we believe there still may be a role for the use of virtual reality in a variety of gynaecological procedures. As the application of VR is inexpensive and usually without adverse events, it is worth trying to optimize its use and study whether it can have a beneficial effect in different populations, different clinical settings or when the application is improved. Therefore we are setting up a study in which we will use VR for the reduction of anxiety and pain during ovarian punctures as part of an IVF treatment, a procedure that is regarded as painful with pain scores varying from 3.9 to 5.6 (scale 1-10) (29). The VR environments available will be critically reviewed and revised to optimize the VR experience.

C. Alternatives to HSG

Our group has recently published the results from the large multicentre randomized controlled trial 'FOAM'. In this study, all 1026 participants are subjected to both an HSG and a HyFoSy with ExEm Foam® in a randomized order. The study compares the results in terms of ongoing pregnancies leading to live births after management after a fertility work-up based on HSG versus management based on HyFoSy. The results of this non-inferiority study show that the procedures result in comparable pregnancy outcomes. Pain scores during HSG were found to be significantly higher in comparison to HyFoSy. Since all participants underwent both procedures, a direct comparison between both tests in terms of their fertility enhancing effect could not be made. For that reason a new study is initiated, that will directly compare HSG with oil contrast versus HyFoSy with ExEm Foam® in terms of live birth within six months after the test. This study is funded by ZonMw (grant number 10390012110083) and is expected to start recruiting in the beginning of 2023 within the Dutch Consortium.

Several groups are working on ways to incorporate tubal flushing with oil-based contrast during or after an alternative tubal patency test such as HyFoSy (30) and THL (31), to combine the advantages of the alternative tubal patency tests with the pregnancy enhancing effect of oil-based contrast.

The results of the diagnostic accuracy review that will be executed, of which the protocol is described in chapter 7, should be taken together with results of therapeutic effects of the various tubal patency tests when deciding on the right test in any given situation. Ideally, patient preference should also be investigated as a factor in deciding what test to advice women in each situation.

D. COVID-19 and infertility

Patients facing infertility experience more psychological distress and feel less able to cope with this distress than patients with endometriosis in our study. A study conducted in America showed that infertility remained the top stressor for infertility patients over other common life stressors as their job, money, health and family. This remained constant before, at the start and amidst the first lockdown (32). In a UK-based study, conducted while gynaecology services and elective surgeries had already restarted after the lockdown, still 55.2% of the 1089 endometriosis patients noticed an increase in their overall pelvic pain and 69.2% reported worsening of associated mental health issues compared to pre-pandemic (33). An international survey study revealed that patients confronted with cancellation or postponement of their medical appointments, fertility treatments or surgery, were more likely to report a deterioration of their mental state (34). This emphasizes the need for a more holistic approach for endometriosis

patients, future research should focus more on ways to give psychological support to both infertility and endometriosis patients.

A research team from our hospital has recently obtained a ZonMw grant (reference number 5160482240003) to investigate the effects of the changes in health care during the pandemic, both in terms of numbers of surgeries and procedures in gynaecology and infertility as well as in terms of patient experiences. This study can aid in better preparation for potential future periods of unavailability of health care.

Future research into eHealth such as telephone- and video appointments should give a closer look into fertility and endometriosis patients to guarantee that the quality of healthcare for these patients does not decrease by innovative and financially attractive interventions. eHealth has been studied in other fields and patient groups over the last decade, and while the results in terms of satisfaction for both the patient and the health care provider are often positive (35, 36), this might not be true for infertility and endometriosis patients.

REFERENCES

1. Zhang J, Lan W, Wang Y, Chen K, Zhang G, Yang W, et al. Ethiodized poppyseed oil-based contrast medium is superior to water-based contrast medium during hysterosalpingography regarding image quality improvement and fertility enhancement: A multicentric, randomized and controlled trial. EClinicalMedicine. 2022;46:101363.
2. Wang R, van Welie N, van Rijswijk J, Johnson NP, Norman RJ, Dreyer K, et al. Effectiveness on fertility outcome of tubal flushing with different contrast media: systematic review and network meta-analysis. Ultrasound Obstet Gynecol. 2019;54(2):172-81.
3. Dreyer K, van Rijswijk J, Mijatovic V, Goddijn M, Verhoeve HR, van Rooij IAJ, et al. Oil-Based or Water-Based Contrast for Hysterosalpingography in Infertile Women. N Engl J Med. 2017;376(21):2043-52.
4. Wang R, Watson A, Johnson N, Cheung K, Fitzgerald C, Mol BWJ, et al. Tubal flushing for subfertility. Cochrane Database Syst Rev. 2020;10:CD003718.
5. Tjon-Kon-Fat RI, Bensdorp AJ, Bossuyt PM, Koks C, Oosterhuis GJ, Hoek A, et al. Is IVF-served two different ways-more cost-effective than IUI with controlled ovarian hyperstimulation? Hum Reprod. 2015;30(10):2331-9.
6. van Eekelen R, van Geloven N, van Wely M, McLernon DJ, Mol F, Custers IM, et al. Is IUI with ovarian stimulation effective in couples with unexplained subfertility? Hum Reprod. 2019;34(1):84-91.
7. Mekaru K, Kamiyama S, Masamoto H, Sakumoto K, Aoki Y. Thyroid function after hysterosalpingography using an oil-soluble iodinated contrast medium. Gynecol Endocrinol. 2008;24(9):498-501.
8. Satoh M, Aso K, Katagiri Y. Thyroid Dysfunction in Neonates Born to Mothers Who Have Undergone Hysterosalpingography Involving an Oil-Soluble Iodinated Contrast Medium. Horm Res Paediatr. 2015;84(6):370-5.
9. van Welie N, Roest I, Portela M, van Rijswijk J, Koks C, Lambalk CB, et al. Thyroid function in neonates conceived after hysterosalpingography with iodinated contrast. Hum Reprod. 2020;35(5):1159-67.
10. Hindocha A, Beere L, O'Flynn H, Watson A, Ahmad G. Pain relief in hysterosalpingography. Cochrane Database Syst Rev. 2015(9):CD006106.
11. Hoffman HG, Chambers GT, Meyer WJ, 3rd, Arceneaux LL, Russell WJ, Seibel EJ, et al. Virtual reality as an adjunctive non-pharmacologic analgesic for acute burn pain during medical procedures. Ann Behav Med. 2011;41(2):183-91.
12. Mallari B, Spaeth EK, Goh H, Boyd BS. Virtual reality as an analgesic for acute and chronic pain in adults: a systematic review and meta-analysis. J Pain Res. 2019;12:2053-85.
13. Aarts JW, van Empel IW, Boivin J, Nelen WL, Kremer JA, Verhaak CM. Relationship between quality of life and distress in infertility: a validation study of the Dutch FertiQoL. Hum Reprod. 2011;26(5):1112-8.
14. Schreurs AMF, van Hoefen Wijsard M, Dancet EAF, Apers S, Kuchenbecker WKH, van de Ven PM, et al. Towards more patient-centred endometriosis care: a cross-sectional survey using the ENDOCARE questionnaire. Hum Reprod Open. 2020;2020(3):hoaa029.
15. van Rijswijk J, van Welie N, Dreyer K, Pham CT, Verhoeve HR, Hoek A, et al. Tubal flushing with oil-based or water-based contrast at hysterosalpingography for infertility: long-term reproductive outcomes of a randomized trial. Fertil Steril. 2020;114(1):155-62.
16. van Welie N, Pham CT, van Rijswijk J, Dreyer K, Verhoeve HR, Hoek A, et al. The long-term costs and effects of tubal flushing with oil-based versus water-based contrast during hysterosalpingography. Reprod Biomed Online. 2021;42(1):150-7.
17. Fertility workup: Dutch Society of Obstetrics and Gynaecology; 2015 [Available from: https://richtlijnendatabase.nl/richtlijn/orienterend_fertiliteitsonderzoek_ofo/startpagina_orienterend_fertilitietsonderzoek.html.

18. Miyamoto Y, Tsujimoto T, Iwai K, Ishida K, Uchimoto R, Miyazawa T, et al. Safety and pharmacokinetics of Iotrolan in hysterosalpingography - Retention and irritability comopared with Lipiodol. Investigative radiology. 1995;30(9):538-43.
19. Mathews DM, Peart JM, Sim RG, Johnson NP, O'Sullivan S, Derraik JGB, et al. The SELFI Study: Iodine excess and thyroid dysfunction in women undergoing oil-soluble contrast hysterosalpingography. J Clin Endocrinol Metab. 2022.
20. So S, Yamaguchi W, Tajima H, Nakayama T, Tamura N, Kanayama N, et al. The effect of oil and water-soluble contrast medium in hysterosalpingography on thyroid function. Gynecol Endocrinol. 2017;33(9):682-5.
21. Patel J, Landers K, Li H, Mortimer RH, Richard K. Thyroid hormones and fetal neurological development. J Endocrinol. 2011;209(1):1-8.
22. Burrow GN, Fisher DA, Larsen PR. Maternal and fetal thyroid function. N Engl J Med. 1994;331(16):1072-8.
23. van Welie N, Portela M, Dreyer K, Schoonmade LJ, van Wely M, Mol BWJ, et al. Iodine contrast prior to or during pregnancy and neonatal thyroid function: a systematic review. Eur J Endocrinol. 2021;184(1):189-98.
24. FDA. FDA recommends thyroid monitoring in babies and young children who receive injections of iodine-containing contrast media for medical imaging 2022 [Available from: https://www.fda.gov/drugs/drug-safety-and-availability/fda-recommends-thyroid-monitoring-babies-and-young-children-who-receive-injections-iodine-containing.
25. Walker MR, Kallingal GJ, Musser JE, Folen R, Stetz MC, Clark JY. Treatment efficacy of virtual reality distraction in the reduction of pain and anxiety during cystoscopy. Mil Med. 2014;179(8):891-6.
26. Fouks Y, Kern G, Cohen A, Reicher L, Shapira Z, Many A, et al. A virtual reality system for pain and anxiety management during outpatient hysteroscopy-A randomized control trial. Eur J Pain. 2022;26(3):600-9.
27. Won AS, Bailey J, Bailenson J, Tataru C, Yoon IA, Golianu B. Immersive Virtual Reality for Pediatric Pain. Children (Basel). 2017;4(7).
28. Deo N, Khan KS, Mak J, Allotey J, Gonzalez Carreras FJ, Fusari G, et al. Virtual reality for acute pain in outpatient hysteroscopy: a randomised controlled trial. BJOG. 2021;128(1):87-95.
29. Buisman E, Roest I, van der Steeg JW, Smeenk JMJ, Koks CAM. Pain scores during oocyte retrieval for IVF/ICSI: A retrospective cohort study comparing three different analgesia protocols. J Gynecol Obstet Hum Reprod. 2022;51(6):102394.
30. De Neubourg D, Janssens L, Verhaegen I, Smits E, Mol BW, Roelant E. Live birth after additional tubal flushing with oil-based contrast versus no additional flushing: a randomised, multicentre, parallel-group pragmatic trial in infertile women with at least one patent tube at hysterosalpingo-foam sonography (HYFOIL study). BMJ Open. 2021;11(11):e054845.
31. Roest I, Hajiyavand AM, Dearn KD, Bongers MY, Mijatovic V, Mol BWJ, et al. Tubal flushing with oil-based contrast during transvaginal hydro laparoscopy, a case report. Facts Views Vis Obgyn. 2022;14:185-8.
32. Vaughan DA, Shah JS, Penzias AS, Domar AD, Toth TL. Infertility remains a top stressor despite the COVID-19 pandemic. Reprod Biomed Online. 2020;41(3):425-7.
33. Spencer JI, Mezquita G, Shakir F. The ongoing impact of the Covid-19 pandemic on endometriosis patients: a survey of 1,089 UK patients. Facts Views Vis Obgyn. 2022;14:257-64.
34. Ashkenazi MS, Huseby OL, Kroken G, Soto-Mota A, Pents M, Loschiavo A, et al. COVID-19 pandemic and the consequential effect on patients with endometriosis. Hum Reprod Open. 2022;2022(2):hoac013.
35. Barsom EZ, Feenstra TM, Bemelman WA, Bonjer JH, Schijven MP. Coping with COVID-19: scaling up virtual care to standard practice. Nat Med. 2020;26(5):632-4.
36. McMaster T, Mori K, Lee S, Manasa S, Stelmach W, To H. Innovations and Implementation of Telemedicine in Surgical Clinics Beyond COVID-19: A Narrative Review. Telemed J E Health. 2022.

APPENDIX

Summary
Samenvatting
List of authors
List of publications
PhD portfolio
Dankwoord
Over de auteur

SUMMARY

Infertility affects approximately one in six couples trying to conceive. Patients facing infertility can be referred for fertility work-up, a series of diagnostic tests, to diagnose and treat or rule out any causes for infertility. If all known cases of infertility are ruled out, this is called unexplained infertility.

One of the tests often performed in the fertility work-up is the hysterosalpingogram (HSG). A contrast fluid is infused into the uterine cavity to delineate the uterine cavity and Fallopian tubes. Once the contrast enters the abdominal cavity through the open ends of the Fallopian tubes, tubal patency is established. Various types of contrast can be used, they can be divided in two groups: oil-based contrast and water-based contrast. Previous research (the H2Oil study) has ascertained that in couples with unexplained infertility and a low risk for tubal pathology, an HSG with use of oil-based contrast results in 10% more ongoing pregnancies than an HSG with use of water-based contrast. It is unclear what the mechanism of action of oil-based contrast is.

Improving pregnancy chances and patient experiences in infertility

In this thesis, we wanted to further investigate the specifics of the fertility-enhancing effect of oil-based contrast during HSG, the safety and accuracy of HSG and test a novel method to reduce discomfort. An additional chapter focusses on the most notable event that took place during the course of this thesis; the Coronavirus-19 pandemic.

Couples with unexplained infertility benefit from an HSG with oil-based contrast, both on the short term and the long term as was shown by a follow-up study using data on pregnancies and fertility treatments up until 5 years after HSG. In **Chapter 2** we conducted a secondary analysis using the database of this long term follow-up study to determine how long the fertility enhancing effect is noticeable. We discovered that the fertility enhancing effect is greatest immediately after HSG, and disappears after around 2 years.

Three important groups of women were excluded from the H2Oil study: women of advanced age, women with ovulation disorders and women with a high risk for tubal pathology based on their medical history. Factors leading to a high risk for tubal pathology are related to having had an pelvic infection, ectopic pregnancy, surgery to the reproductive organs and/or endometriosis. We are repeating the H2Oil study but this time including exactly these groups of women. The protocol of this ongoing study, called ‘H2Oil2’, is presented in **Chapter 3**. We expect to use the results of this study to

renew the infertility guidelines with evidence-based advice on the type of contrast to use in the three studied patient groups.

The database of the first H2Oil study comprises 1119 women with unexplained infertility and details of their fertility work-up, treatments and outcomes are all reported. This present a unique opportunity to replicate a previously formulated hypothesis. In previous work, the prognosis for natural conception (poor, medium or good) was found to be a predictor for the benefit of intra-uterine insemination with mild ovarian stimulation (IUI-MOH) treatment as opposed to expectant management. After performing the same analysis on the H2Oil database, in **Chapter 4**, this interaction between prognosis and therapeutic effect of IUI-MOH was not found. We did, however, confirm that couples with a longer duration of infertility benefitted more from IUI-MOH in comparison to expectant management. These findings can also be used by physicians to advice couples with a short duration of infertility to pursue natural conception.

Despite the convincing results of the H2Oil study which were published in 2017, we noticed that physicians were still hesitant to use oil-based contrast. One of the reasons was a fear of complications. When the contrast fluid enters the bloodstream through uterine vessels, this so-called intravasation can lead to oil-embolism and can have severe complications when an oil-embolus reaches the lungs or brain circulation. In **Chapter 5** we performed a systematic literature review to obtain information on all possible complications of HSGs with oil-based contrast. We did not use a time restriction or language restriction to make sure that all useful and available articles were found. Combining all published data from 8 RCTs, 41 cohort studies and 59 case reports, the incidence of intravasation was found to be 2.7% during HSGs with oil-based contrast, in comparison to 2.0% with water-based contrast. The second most frequently reported complication was infection, which was reported in 20 studies. The frequency was 0.9% after the use of oil-based contrast and 1.9% after water-based contrast, however, when including only studies published after the widespread use of antibiotics these numbers decrease to 0.55% and 0.35% respectively. Based on these results we conclude that safety concerns should not be a reason to withhold women an HSG with oil-based contrast.

The HSG is an uncomfortable or painful procedure for most women, with studies reporting a pain score of 3 to 5 on a scale of 0-10. Multiple types of pain medication have been tried, including tablets, intravenous medication and local injections. None of these give a satisfactory reduction in pain without major side effects. In **Chapter 6** we described our randomized controlled trial in which we studied whether the use of Virtual Reality can reduce discomfort and/or pain. Virtual Reality is a relatively new technique that has proven effective in pain reduction during various medical procedures. As it

distracts the user from their surroundings and takes them into a virtual world, the brain is focussed on the virtual reality instead of bodily sensations such as pain. Our study did not show a significant difference in overall pain, peak pain or satisfaction between the Virtual Reality group and the control group. It is possible that the VR intervention was not as effective because of a relatively low level of 'immersion', the feeling of being present in the VR environment. This low level of immersion in our study can be explained by the lack of sound in addition to the glasses and the fact that the movies were a lot shorter than the procedure, meaning that the participant repeatedly had to choose a new movie.

HSG was introduced as a diagnostic test. The safety and comfort issues mentioned, but also radiation exposure, the use of iodine containing contrast and the need for a radiologic department were all reasons for the development of other types of visual tubal patency tests. These tests are based on different modalities such as ultrasound (Hysterosalpingo-foam-sonography or HyFoSy, Hysterosalpingo-contrast-sonography or HyCoSy, both either in 2D, 3D or 4D), MRI (MR-HSG) and transvaginal hydrolaparoscopy (THL). These alternative tests have all been studied with regards to efficacy, safety and diagnostic quality, but a review comparing the different tests with laparoscopy, the current gold standard, is missing. This is the reason for our Cochrane Review for which the protocol is presented in **Chapter 7**.

During the course of this PhD-tract, Coronavirus Disease 2019 became a pandemic in the spring of 2020. All non-essential healthcare, including fertility care, was paused to prevent further spreading of the disease and to use all healthcare personnel and resources for the diseased. For patients with infertility this was a very uncertain time. Scheduled appointments were converted to telephone or video appointments and new treatment cycles could not be started. In **Chapter 8** we described the results of our national survey, showing a low fertility-related quality of life in women with infertility during the COVID-19 pandemic and lockdown. Patients gave a good rating to the telephone and video appointments as a replacement of regular consultations during the lockdown, and thought that regular telephone appointments were a good addition to regular care in the future.

NEDERLANDSE SAMENVATTING

Ongeveer 1 op de 6 koppels met een kinderwens krijgt te maken met vruchtbaarheidsproblemen. Deze koppels kunnen verwezen worden voor het oriënterend fertiliteitsonderzoek, een serie aan vragen en testen om de verschillende oorzaken voor vruchtbaarheidsproblemen aan te tonen of uit te sluiten. Als er geen oorzaak wordt gevonden spreken we van onverklaarde subfertiliteit.

Eén van de testen uit het fertiliteitsonderzoek is het hysterosalpingogram (HSG). Hierbij wordt contrastvloeistof in de baarmoeder gebracht zodat de baarmoederholte en de eileiders in beeld kunnen worden gebracht met röntgenfoto's. Zodra de contrastvloeistof uit de eileiders in de buikholte loopt bevestigd dit dat de eileiders open zijn. Er zijn verschillende soorten contrastmiddelen die gebruikt kunnen worden. De twee grootste groepen zijn oliehoudend contrast en waterhoudend contrast. Eerder onderzoek (de H2Olie studie) liet zien dat in koppels met onverklaarde subfertiliteit en een laag risico op eileiderverstopping, een HSG met oliehoudend contrast tot 10% meer zwangerschappen leidt dan een HSG met waterhoudend contrast. Het is nog onduidelijk hoe dit therapeutische effect precies werkt.

Het verbeteren van de zwangerschapskansen en ervaringen van subfertiele patiënten

In deze dissertatie was ons doel om meer duidelijkheid te krijgen over het therapeutische effect van oliehoudend contrast tijdens het HSG, de veiligheid en accuraatheid van het HSG en we onderzochten een nieuwe methode om het HSG minder onaangenaam te maken. Een extra hoofdstuk in deze dissertatie heeft te maken met de grootste gebeurtenis van de afgelopen jaren: de COVID-19 pandemie.

Koppels met onverklaarde subfertiliteit hebben baat bij een HSG met oliehoudend contrast, zowel op korte als lange termijn. Dit werd aangetoond in een studie die tot 5 jaar na het HSG gegevens verzamelde. In **Hoofdstuk 2** hebben we een secundaire analyse uitgevoerd in deze lange termijn-database. Hiermee konden we onderzoeken wat de duur van het therapeutische effect is. We kwamen tot de conclusie dat het positieve effect van oliehoudend contrast het grootst is direct na het HSG, en daarna langzaam af neemt tot 2 jaar na het HSG.

Drie grote groepen vrouwen werden niet meegenomen in de H2Olie studie: vrouwen met een hogere leeftijd, vrouwen met ovulatieproblemen en vrouwen die een hoog risico hebben op eileiderproblemen op basis van hun medische voorgeschiedenis. Factoren die bijdragen aan een hoog risico op eileiderproblemen zijn: een infectie in de onderbuik,

een buiten baarmoederlijke zwangerschap, een operatie aan de eierstokken of eileiders en endometriose. Om het therapeutische effect ook in deze groepen vrouwen te onderzoeken herhalen we de H2Olie studie en includeren we nu júist deze drie groepen vrouwen. Deze studie heet 'H2Olie2' en het protocol wordt beschreven in **Hoofdstuk 3**. De studie loopt nog, we verwachten dat de resultaten van deze studie kunnen bijdragen aan herziening van de landelijke richtlijnen voor gynaecologen.

In de H2Olie studie van 2017 zaten 1119 vrouwen met onverklaarde subfertiliteit, waarvan alle gegevens van hun onderzoeken en behandelingen bekend zijn. Daarmee is dit een waardevolle database met (geanonimiseerde) patiëntgegevens. We hebben deze database gebruikt om een eerdere studie te herhalen en zo de uitkomsten te vergelijken. In de eerdere studie werd gezien dat de prognose voor natuurlijke zwangerschap (slecht, gemiddeld of goed) van invloed was op het effect van inseminatiebehandelingen ten opzichte van een afwachtend beleid. Bij het herhalen van de analyses in de H2Olie database, in **Hoofdstuk 4**, zagen we deze relatie niet terug. We konden wel bevestigen dat koppels met een langere duur van subfertiliteit méér baat hadden bij inseminaties dan koppels die pas kort subfertiel waren. Deze uitkomst kan door artsen gebruikt worden om koppels met een korte duur van onverklaarde subfertiliteit te adviseren om nog niet direct te starten met een behandeling.

Ondanks de overtuigende resultaten van de H2Olie studie uit 2017 merkten we dat sommige artsen nog steeds huiverig zijn om oliehoudend contrast te gebruiken. Eén van de redenen hiervoor is de angst voor complicaties. Het oliehoudende contrastvloeistof in de baarmoederholte kan soms de bloedvaten van de baarmoeder binnendringen, dit noemen we intravasatie. Als de contrastvloeistof oliedruppels vormt in de bloedvaten en naar de longen of hersenen stroomt kan dit ernstige gevolgen hebben. In **Hoofdstuk 5** hebben we een systematische literatuurstudie uitgevoerd om informatie over alle soorten complicaties van een HSG met oliehoudend contrast te verkrijgen. We hebben studies uit alle jaren en in alle talen verzameld zodat er geen informatie verloren ging. Nadat we alle gepubliceerde gegevens van 8 gerandomiseerde klinische studies, 41 cohort studies en 59 patiëntbeschrijvingen bij elkaar hebben genomen is het percentage van patiënten met intravasatie 2.7% na een HSG met oliehoudend contrast en 2.0% na een HSG met waterhoudend contrast. De tweede meest genoemde complicatie was infectie, dit werd beschreven in 20 studies. De frequentie van een infectie was 0.9% na een HSG met oliehoudend contrast en 1.9% na een HSG met waterhoudend contrast. Als we de studies weglaten die rapporteren over gegevens van vóór het gebruik van antibiotica, dalen deze cijfers naar 0.55% voor oliehoudend contrast en 0.35% voor waterhoudend contrast.

Het HSG is voor de meeste vrouwen een oncomfortabele of pijnlijke procedure. De gemiddelde pijnscore is 3 tot 5 op een schaal van 0 tot 10. Verschillende soorten pijnmedicatie zijn geprobeerd, zoals tabletten, injecties en medicatie via het infuus. Geen van deze opties geeft voldoende resultaat zonder nadelige effecten. In **Hoofdstuk 6** beschrijven we de resultaten van onze gerandomiseerde studie, waarin we onderzoeken of *Virtual Reality* kan helpen de pijn te verminderen. *Virtual Reality* is een relatief nieuwe techniek in de medische wereld die bewezen effectief is om pijn te verminderen. Het neemt de gebruiker mee in een virtuele wereld en daardoor worden de hersenen afgeleid van lichamelijke sensaties als pijn. Onze studie naar het effect van VR tijdens het HSG liet geen effect zien in totale pijnscore, piek-pijnscore of tevredenheid tussen de groep vrouwen die VR kreeg tijdens het HSG en de controlegroep zonder VR tijdens het HSG. Mogelijk kan dit gebrek aan effect verklaard worden doordat de VR in onze studie niet heel hoog scoorde in meeslependheid en het aanwezig zijn 'in' de andere realiteit. In onze studie was er geen geluid bij de films op de VR-bril en de films waren korter dan de HSG procedure, waardoor er steeds een korte onderbreking was in de afleiding.

Het HSG is initieel ontworpen als diagnostische test. Naast de zorgen om complicaties en de pijn zijn er nog andere nadelen aan het HSG: er wordt gebruik gemaakt van röntgenstraling en jodiumhoudend contrast en er is een afdeling radiologie nodig. Daarom zijn er andere soorten testen ontwikkeld die ook naar de eileiders kijken. Hierbij wordt gebruik gemaakt van echo (zoals de HyFoSy en HyCoSy echo), MRI (MR-HSG) en een vaginale kijkoperatie (THL). Deze alternatieve testen zijn allemaal los van elkaar onderzocht op effectiviteit, veiligheid en diagnostische kwaliteit maar ze zijn nooit goed met elkaar vergeleken. Dat willen we in de studie uit **Hoofdstuk 7** gaan doen.

Gedurende de looptijd van de bovenstaande hoofdstukken brak in het voorjaar van 2020 COVID-19 uit als pandemie. Alle niet-essentiële zorg, waaronder zorg voor koppels met subfertiliteit, kwam stil te liggen zodat alle aandacht naar de bestrijding van de pandemie kon gaan. Voor subfertiele patiënten was dit een heel onzekere tijd en werden veel afspraken afgezegd of omgezet naar telefonische afspraken. Twee maanden lang werden geen nieuwe vruchtbaarheidsbehandelingen opgestart. In **Hoofdstuk 8** beschrijven we de resultaten van onze landelijke vragenlijststudie. Hierin laten we zien dat de vruchtbaarheids-gerelateerde kwaliteit van leven lager was in subfertiele vrouwen gedurende de lockdown. Patiënten vonden het fijn dat afspraken werden omgezet naar telefonische afspraken en video-afspraken, en ze zouden graag meer telefonische afspraken krijgen als aanvulling op reguliere zorg in de toekomst.

LIST OF AUTHORS

Amsterdam UMC, Academic Medical Center, Amsterdam, The Netherlands, Department of Epidemiology and Data Science
Rik van Eekelen

Amsterdam UMC, Academic Medical Center, Amsterdam, The Netherlands, Department of Obstetrics and Gynaecology
Inge M Custers
Rachel Tros

Amsterdam UMC, Academic Medical Center, Amsterdam, The Netherlands, Department of Reproductive Medicine
Mariëtte Goddijn
Madelon van Wely

Amsterdam UMC, Vrije Universiteit, Amsterdam, The Netherlands, Department of Radiology and Nuclear Medicine
Jaap Stoker

Amsterdam UMC, Vrije Universiteit, Amsterdam, The Netherlands, Department of Reproductive Medicine
Jasmijn Bergwerff
Kim Dreyer
Annelotte P van Haaps
Danah Kamphuis
Cornelis B Lambalk
Laura L van Loendersloot
Velja Mijatovic
Joukje van Rijswijk
Anneke MF Schreurs
Nienke van Welie

Catharina Ziekenhuis, Eindhoven, The Netherlands, Department of Obstetrics and Gynaecology
Dana AG Huppelschoten

De Endometriose Stichting, Sittard, The Netherlands
Bianca De Bie

Elkerliek Ziekenhuis, Helmond, The Netherlands, Department of Obstetrics and Gynaecology
Tine van de Laar

Freya, Gorinchem, The Netherlands
José Knijnenburg

Imperial College Healthcare NHS Trust, London, England, Department of Radiology
Anne Hemingway

Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands
Marinus JC Eijkemans

Leiden University Medical Center, Leiden, the Netherlands, Department of Biomedical Data Sciences
Nan van Geloven

Maastricht UMC+, Maastricht, The Netherlands, Department of Gynaecology
Jacques WM Maas

Máxima Medisch Centrum, Veldhoven, Netherlands, Department of Obstetrics and Gynaecology
Marlies Y Bongers
Carolien Koks
Inez Roest

Monash University, Clayton, Australia, Department of Obstetrics and Gynaecology
Wentao Li
Ben WJ Mol
Rui Wang

Radboudumc, Nijmegen, The Netherlands, Department of Gynaecology
Annemiek W Nap

Reinier de Graaf Gasthuis, Delft, The Netherlands, Department of Reproductive Medicine
Evert JP van Santbrink

Zaans Medisch Centrum, Zaandam, The Netherlands, Department of Obstetrics and Gynaecology
Angelo B Hooker

Ziekenhuis Amstelland, Amstelveen, The Netherlands, Department of Obstetrics and Gynaecology
Alexander Mozes

LIST OF PUBLICATIONS

** De publicaties die onderdeel uitmaken van dit proefschrift zijn aangegeven met een asterix*
‡ Gedeeld eerste auteurschap is aangegeven met een dubbel kruis

*2020, Hum Rep Open	Does the effectiveness of IUI in couples with unexplained subfertility depend on their prognosis of natural conception? A replication of the H2Oil study *Van Eekelen R, **Rosielle K**, van Welie N, Dreyer K, Mol BWJ, Eijkemans MJ, Mijatovic V, van Geloven N*
*2020, RBMO	How long does the fertility-enhancing effect of hysterosalpingography with oil-based contrast last? ***Rosielle K**‡, Van Welie N‡, van Rijswijk J, Lambalk CB, van Geloven N, Mijatovic V, Mol BWJ, van Eekelen R*
*2021, RBMO	Safety of oil-based contrast medium for hysterosalpingography: a systematic review *Roest I, **Rosielle K**, van Welie N, Dreyer K, Bongers M, Mijatovic V, Mol BWJ, Koks C*
*2021, RBMO	The impact of the COVID-19 pandemic on infertility patients and endometriosis patient in the Netherlands ***Rosielle K**‡, Bergwerff J‡, Schreurs AMF, Knijnenburg J, De Bie B, Maas JWM, Nap AW, van Wely M, Lambalk CB, Goddijn M, Custers IM, van Loendersloot LL, Mijatovic V*
2021, NTOG	Het effect van de COVID-19 pandemie op fertiliteitspatiënten en endometriose patiënten ***Rosielle K**, Bergwerff J, Schreurs AMF, Knijnenburg J, De Bie B, Maas JWM, Nap AW, van Wely M, Lambalk CB, Goddijn M, Custers IM, van Loendersloot LL, Mijatovic V*
2022, NTOG	De tuba in (onverklaarde) subfertiliteit *Kamphuis D, van Welie N, **Rosielle K**, Roest I, van Rijswijk J, van Wely M, Bossuyt PMM, Koks CAM, Finken MJJ, Bongers MY, Lambalk CB, Mol BWJ, Dreyer K, Mijatovic V*

*2022, BMC Women's Health	Oil-based versus water-based contrast media for hysterosalpingography in infertile women of advanced age, with ovulation disorders or a high risk for tubal pathology: study protocol of a randomized controlled trial (H2Oil2 study) ***Rosielle K**, Kamphuis D, van Welie N, Roest I, Mozes A, van Santbrink EJP, van de Laar T, Hooker AB, Huppelschoten AG, Li W, Bongers MY, Stoker J, van Wely M, Koks C, Lambalk CB, Hemingway A, Mol BWJ, Dreyer K, Mijatovic V*
*2022, Cochrane database of systematic reviews	Visual tubal patency tests for tubal occlusion and hydrosalpinx *Tros R, **Rosielle K**, Koks C, Mijatovic V, Bongers MY, Mol BWJ, Wang R*
2023, BMC Women's Health	The effectiveness of immediate versus delayed tubal flushing with oil-based contrast in women with unexplained infertility (H2Oil-timing study): study protocol of a randomized controlled trial. *Kamphuis D, **Rosielle K**, van Welie N, Roest I, van Dongen AJCM, Brinkhuis EA, Bourdrez P, Mozes A, Verhoeve HR, van der Ham DP, Vrouenraets FPJM, Risseeuw JJ, van de Laar T, Janse F, den Hartog JE, de Hundt M, Hooker AB, Huppelschoten AG, Pieterse QD, Bongers MY, Stoker J, Koks CAM, Lambalk CB, Hemingway A, Li W, Mol BWJ, Dreyer K, Mijatovic V*

PHD PORTFOLIO

PhD Student:	Kimmy Rosielle
PhD period:	March 2019 – December 2022
PhD supervisors:	Prof. Dr. V. Mijatovic, Prof. Dr. C.B. Lambalk, Dr. K. Dreyer, Prof. Dr. B.W.J. Mol

Presentations	
2023	Presentation 'Hysterosalpingografie – geschiedenis, effect en complicaties' Future of Medical Imaging and Radiology 2023, Almere
2020	Presentation 'The effect of the COVOID-19 pandemic on women facing infertility in the Netherlands' (*Held by dr. van Loendersloot)* ESHRE 2022 (digital)
2020	Presentation 'Patiëntervaringen met endometriose- en fertiliteitszorg tijdens de intelligente lockdown' Refeeravond Amsterdam UMC (digital) V&VN symposium (digital)
Conference attendance	
2022	ESRHE Milan
2020	ESHRE Online
2020	UK Fertility
Grant	
2020	ZonMw grant in program 'COVID-19, wetenschap voor de praktijk' Title project: The impact and consequences of the restrictive measures following the COVID-19 outbreak for fertility patients and patients with endometriosis in the Netherlands.
Teaching	
2022	Supervision high school students with research internship, 6 months
2020	Supervision master student Medicine with research internship, 4 months
2019 – 2022	Pre internship training Gynaecology for Masters' students Medicine

Courses	
2022	Basic (Medical) Statistics, Amsterdam UMC Doctoral School
2021	Writing a Scientific Article, Taalcentrum VU
2020	Scientific Integrity, VUmc Academy
2019, 2020 & 2021	AR&D Retreat
2019	Data Management plan, UBVU
2019	Castor EDC Training, Clinical Research Unit AMC
2019	BROK, NFU

DANKWOORD

Het boek is af! Wat een groot en mooi project, en wat is het toch ook wel fijn dat er met een proefschrift en verdediging een einde aan komt aan deze periode. De afgelopen jaren heb ik veel geleerd over het opzetten, uitvoeren en interpreteren van onderzoek. Ik heb met veel mensen om mij heen samengewerkt en van iedereen veel geleerd, een aantal mensen wil ik graag specifiek bedanken.

Allereerst wil ik de vrouwen bedanken die hebben deelgenomen aan de diverse klinische studies, waarvan er een aantal in dit proefschrift beschreven staan. Door hun deelname hebben we meer inzicht verkregen in de fertiliteitsonderzoeken en fertiliteitszorg en zo kunnen we vrouwen in de toekomst hopelijk beter helpen.

Geachte leescommissie, *prof. dr. Oudijk, prof. dr. Hoek, prof. dr. Laven, prof. dr. de Neubourg, dr. Kallewaard en dr. van Waesberghe*, hartelijk dank voor de tijd die jullie genomen hebben om mijn proefschrift zorgvuldig en kritisch te lezen.

Prof. dr. Mijatovic, beste *Velja*, ik wil je bedanken voor de leerzame tijd. Je bent betrokken, begripvol, je hebt voor ieder probleem een oplossing of kent iemand die kan helpen. Altijd een positieve insteek. Je bent voor de onderzoekers gemakkelijk bereikbaar, zelfs nu je poli draait op locatie AMC en we je niet meer zo makkelijk 'aan je jasje kunnen trekken' en dat bewonder ik! Naast het inhoudelijke heb ik ook veel van je geleerd op het gebied van samenwerking en communicatie met verschillende partijen. Dank voor je vertrouwen en je begeleiding!

*Prof. dr. Lambalk***,** beste *Nils***,** dank voor je wijze lessen en kritische houding. Overvol met wilde wetenschappelijke theorieën, te veel onderzoeksideeën om in één carrière uit te werken en vaak weer een verrassende invalshoek of tip voor het verbeteren van een manuscript. Ik vind het jammer voor de nieuwere onderzoekers dat ze je mini-colleges tijdens de besprekingen moeten missen, ik heb ontzettend veel van je geleerd.

Dr. Dreyer, beste *Kim*, wat was ik blij om jou als copromotor te hebben om af en toe de vertaalslag te kunnen maken van mij als jonge onderzoeker naar de professoren. Met jouw ervaring van het uitvoeren van onder andere de H2Olie-studie kon je me goed helpen bij de nieuwe studies, je wist wat de valkuilen waren en hoe ik logistieke en inhoudelijke problemen het beste kon aanpakken. Bewonderenswaardig hoe je altijd nauwkeurig en kritisch naar teksten kunt blijven kijken ook als je ze al 10x onder ogen hebt gehad. Bedankt!

Prof. dr. Mol, beste *BW*, vanaf de andere kant van de wereld blijf je betrokken bij de onderzoeksgroep. Indrukwekkend hoe je van zo'n afstand mensen weet te verbinden en altijd een advies hebt wanneer nodig. Je hebt me geholpen samenwerkingen aan te gaan met verschillende andere onderzoekers en ik heb van je geleerd af en toe uit te zoomen om de juiste richting te kiezen en focus te verleggen. Dank hiervoor!

COLLEGA'S

Nienke, door jou harde werk en grondige voorbereidingen kon ik met vliegende start beginnen aan mijn promotietraject. Juist dankzij onze verschillen heb ik ontzettend veel van je geleerd. *Danah*, jij hebt de H2Olie-studies van me overgenomen en sindsdien zijn ze bij jou in goede handen. Fantastisch dat de H2Olie-timing studie inmiddels vol zit. Ik heb genoten van onze gesprekken over olie en over het leven! Beste *Inez,* ons H2Olie-teamlid in het Zuiden. Dank voor de fijne samenwerking bij het review-artikel en je inzet voor de H2Olie-studies in de zuidelijke provincies.

Annelotte, jaren een heel gezellig kamergenootje en fietsbuddy, maar bovenal mijn VR-vriendin! Wat begon als ondersteuning tijdens mijn verlof werd een volledige samenwerking in de VICTORY en je bent druk met VR-vervolg-studies. Top om te zien hoe je alles oppakt! Ik mis je bakkunsten wel…

Lieve *Daan* en *Lot,* het 'verdwijnen' van de poli, de verhuizing van -1 naar 8D: we hebben wat meegemaakt met elkaar als dagelijkse VPG-club! Ik heb genoten van de cappu's (graag 'lekker groot') en de vele ESHRE-outfit-besprekingen.

Mede-VPG-er *Anneke,* altijd was je bereid te helpen met ABR-formulieren, METc-indieningen en ander administratief / bureaucratisch gedoe. Daarnaast heb je eigenhandig de éne na de andere onderzoeker geïnspireerd met je duurzame lifestyle! Fijn om van je te kunnen leren met ons gezamenlijke COVID-onderzoek. *Marissa,* geen VPG-onderzoeker, wel mijn kamergenootje van begin tot eind. Verfrissend om ook een niet-VPG-er in ons midden te hebben en je bent altijd een goede sparringpartner. We zijn dit jaar tegelijk gestart met de opleiding en ik heb zin om de gezelligheid door te zetten bij alle komende cursussen!

Rui & Rachel, I will start with apologies for missing those Cochrane-study-meetings in the first few months of my residency.. We're currently powering through >10.000 studies to screen, I know we will make it thanks to Rui's unlimited trust and uplifting emails!

Prof. dr. Hemingway, dear *Anne,* I have really enjoyed working with you and your wonderful team including *Shaz, Lesley and Katherine.* Hopefully we'll find a reason to meet up again shortly. Thank you for your enthusiasm!

Lieve *medewerkers van de (voormalige) poli-VPG en medewerkers van het IVF-centrum:* Dank voor al jullie hulp, gezelligheid, het uitdelen van patiënten informatie over de studies en natuurlijk de snoepjes. *Yvonne*, jij kon me altijd helpen bij 'hoe dingen werken op de poli' en ik vind het jammer dat ik door COVID niet vaker heb kunnen genieten van je kookkunsten bij jaarlijkse VPG-uitjes. *Madeleine*, je werkte als secretaresse van de VPG maar hielp bij veel meer! Door het vele thuiswerken kon ik niet meer zo gemakkelijk bij je 'binnenlopen' maar je was altijd bereid te helpen. Veel succes met je nieuwe uitdaging bij de METc!

Lieve *OVGO-meiden*! Mede door de gezelligheid van jullie allen voelde 'de VU' echt als een warm bad. *Carry en Emilie,* we leerde elkaar kennen in het OLVG Oost en kwamen één voor één naar de VU voor onderzoek, hoog tijd om weer eens te glitteren! *Rianne*, naast OVGO-collega ook de gezelligste buurvrouw op de Vrolik, ik heb genoten van de borrels. Wat fijn *Joyce* dat jij altijd in was voor een (oriënterend) loopje naar de AH om het leven te bespreken. Mei, Anna, Asra, Tessa, Jill, Feline, Saskia, Lisa, Emmy.. En hiermee heb ik nog lang niet iedereen benoemd… Ik ben blij dat ik vóór COVID nog heb kunnen genieten van alle leuke etentjes, lunches, het zeilweekend, de wintersport en ga zo maar door. Ik hoop en verwacht velen van jullie tegen te blijven komen binnen de gynaecologie!

Bestuur *ProVUmc* en later *ASAP*, wat was het leuk en goed om met jullie samen te werken! APROVE-ers, bedankt voor de fijne samenwerking. Top dat we in vrij korte tijd de twee verenigingen tot ASAP hebben kunnen bouwen waarbij we alle 3000 (!!) promovendi van AmsterdamUMC vertegenwoordigen!

Dank ook aan alle *OLVG-ers* waarmee ik heb samengewerkt en die mij hebben geholpen bij mijn eerste stappen in de gynaecologie en verloskunde. Onzeker of die gynaecologenwereld wel bij mij paste begon ik bij jullie met werken, en gedurende de drukke avonddiensten, poli-dagen en borrelavonden bij Bukowski kwam ik er achter dat de passie blijvend én de moeite waard was. In het bijzonder dank aan *Eugenie* als opleider en *Erica* als mentor, ik heb heel veel gehad aan onze gesprekken destijds en meer recent voor mijn sollicitatie voor de opleiding tot gynaecoloog.

Collega's in het *Tergooi ziekenhuis*, wat fijn dat jullie me met open armen hebben ontvangen. Ik weet zeker dat ik heel veel van jullie ga leren de komende jaren!

PARANYMFEN

Mijn geneeskunde-buddy's van het eerste uur! Lieve *Kyr* en *Nouk,* vanaf de eerste treinritjes Driebergen-Bijlmer hebben we elkaar door de studie heen geholpen en toen ik begon met dit promotietraject zaten jullie vol tips en waarschuwingen door jullie eigen ervaringen. De 3-maanden-dip, de frustraties en de berg op 't einde; door jullie wist ik dat dit er bij hoorde en na een middag thee, avondje wijn of nacht dansen samen voelt alles altijd weer een stukje lichter. Ik ben ontzettend trots dat jullie aan mijn zijde staan!

FAMILIE EN VRIENDEN

Bregje en *Kris,* ook al wonen we ver uit elkaar en zijn onze levens zo verschillend, het is altijd weer alsof niets is veranderd sinds 't Went.. *Esmee*, van groepje 51, naar gyn-kind-collega's in het OLVG en nu weer in Tergooi, ik hoop dat we elkaar blijven opzoeken, in de speeltuin of in het ziekenhuis. Ook de andere *51-ers* bedankt voor jullie gezelligheid! Lieve *Ann* en *Mary,* wat een jaar is 2023 voor ons drieën. Sinds de middelbare school delen we werkelijk alles samen, ik hoop dat ik straks weer ietsje meer tijd heb voor shop/sauna/dans/wat-dan-ook-activiteiten met jullie. Dank voor jullie steun en liefde al die jaren! Lieve familie en in het bijzonder *Robin, Eva, Ilona, Terry en Hennie*, bedankt voor de gezelligheid! Lieve schoonfamilie, bedankt voor jullie interesse en hulp, ik voel me al meer dan 10 jaar heel welkom.

Lieve *Iraj* en *Nasrin,* bedankt voor jullie steun, gesprekken en advies in de afgelopen jaren. *Sawan*, door jou te zien opgroeien leer ik ook steeds meer over mijzelf, bedankt daarvoor!

Lieve *mama en Harm,* bedankt voor jullie liefde en hulp. *Mam*, je hebt me altijd onvoorwaardelijk gesteund en de vrijheid en veiligheid gegeven om me te ontwikkelen, dat waardeer ik enorm! Harm, inmiddels echt familie, bedankt voor je gezelligheid en steun.

Liefste *Martijn*, inmiddels 5 jaar getrouwd, ik kan me geen leven voorstellen zonder jou. Je steunt me in alles wat ik doe en helpt me te relativeren wanneer ik dat nodig heb. Alles is fijner en gezelliger samen! Na de relatief rustige jaren tijdens mijn promotie ben ik nu weer drukker met werk en diensten, samen komen we er gelukkig altijd uit. We leerden elkaar kennen hier in Amsterdam (Diemen..) en ik kijk er enorm naar uit om samen ons leven verder op te bouwen in Amersfoort. Lieve *Milo*, onze kleine *Maki*, jij bent ons grootste geluk en ik geniet er elke dag van de wereld samen met jou te ontdekken!

ABOUT THE AUTHOR

Kimmy Rosielle was born on April 20th 1991 and grew up in Doorn with her mother. After graduating high school at Revius Lyceum in Doorn in 2009, she started studying Biology at Utrecht University. In 2010 she switched to study Medicine at the University of Amsterdam and moved to Diemen. Eager to start clinical work, she did an extracurricular internship at the department of Urology at Kuopio hospital in Finland after finishing her Bachelor's degree. She focused her Master's thesis on the effect of hyperbaric treatment for haemorrhagic radiation cystitis, at the department of hyperbaric medicine in the Academic Medical Center. Her interest for Obstetrics and Gynaecology was first sparked during her clinical internship at the OLVG in Amsterdam. She finished her Master's degree with a final internship in the same department and obtained her first job there after graduating in 2017. She worked at OLVG Amsterdam from 2017 until 2019, as she started as a PhD student at the Amsterdam UMC, Vrije Universiteit. Her research focussed on the therapeutic effects of hysterosalpingography in women with infertility. During her time as a PhD student she was board member of the PhD-student council ProVUmc and later ASAP. In 2023 Kimmy started her residency in Obstetrics and Gynaecology at Tergooi Ziekenhuis, as part of her training to become a gynaecologist at Amsterdam UMC.

Kimmy got married to Martijn in 2018 and their son Milo was born in 2021.