REPRODUCTIVE OUTCOMES IN WOMEN WITH LOW OVARIAN RESERVE

By

DR BALA MURUHAN KARUNAKARAN

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Institute of Metabolism and Systems Research

College of Medical and Dental Sciences

University of Birmingham

February 2019

Abstract

The number of women with low ovarian reserve seeking fertility treatment is increasing, due to advancing maternal age at conception. Women with low ovarian reserve have a low IVF success rate. This thesis aims to increase our understanding of women with low ovarian reserve, their reproductive outcomes and their reproductive physiology. The evidence is synthesised using two systematic reviews, a prospective cohort study, a retrospective analysis of data and two qualitative studies. The main findings are:

1. Low ovarian reserve, quantified by AFC, AMH and FSH, is associated with low live birth rates and incidences of pregnancy loss after assisted reproduction.
2. There is inter-cycle variation in AFC, AMH and FSH in women. In this cohort, FSH and AFC appear to have a higher magnitude of variation in comparison to AMH.
3. There is inter-cycle variation in AFC, AMH and FSH in women with low ovarian reserve.
4. Clinicians find treating women with low ovarian reserve challenging. Women with low ovarian reserve are unaware of their low IVF success rates and there is cultural and religious stigma about the acceptance of egg donation. Both clinicians and women with low ovarian reserve express willingness to take part and support research studies.

Dedication

This thesis is dedicated to my parents, Dr Velupillai Karunakaran and Dr Vasantha Maligah Karunaharan. Without their countless sacrifices and unconditional love, I would not have been able to achieve a fraction of what I have today.

My father, Dr Velupillai Karunakaran, a retired obstetrician and gynaecologist, was very proud of me securing my PhD studentship. I sorely miss him and wish he is still with us to see me reach the finishing line.

Acknowledgements

"If I have seen a little further, it is by standing on the shoulders of giants."

Sir Isaac Newton, 1676

My PhD has been a journey with multiple challenges on the way. Without the support of many, it would not have been possible to complete this.

I would like to thank my mentor and primary supervisor Professor Arri Coomarasamy, for taking me on in his team and offering guidance and support.

I would like to thank my second supervisor, Dr Ioannis Gallos, for his methodology guidance and lending his statistical expertise.

I thank my second reviewers Dr Rima Smith and Dr Abey Eapen, as well as my second coder Dr James Cheshire. I would also like to thank Dr Abey Eapen for teaching me AFC and helping me with my scanning. I thank Miss Emily Colley for chaperoning most of my scans, outside normal working hours.

I thank Mr Aurelio Tobias for his statistical guidance. I thank Mr Adam Devall and Miss Helen Williams for helping me navigate all the required administration to get my project approved. I thank Mr Yusuf Muhammed for his help with teaching me to create some of the illustrations in this thesis. I thank Dr Justin Chu for providing his study consent form and patient information sheets for me to use as templates. I thank Derick Yates and Jon Andrews for their help with formulating search strategies for the systematic reviews.

I thank Dr Rebecca Man for her continuous moral support, encouragement and comments on this thesis. I thank Dr Melody Redman, Dr Tina Verghese, Dr

Catherine Dunlop, Dr Amie Wilson, Dr Donald Masi and Dr Laurentiu Craciunas for their constructive comments on various chapters of this thesis. I would like to thank all the clinical research fellows and lecturers, collectively known as the United Colours of Benetton, at the BWCH for their comradery, support and being a great sounding board for ideas.

I would like to thank all the women and clinicians who willingly gave their time altruistically for this project.

Table of contents

Contents

List of illustrations

List of tables

Abbreviations

ACU	Assisted conception unit
AFC	Antral follicle count
AMH	Anti-mullerian hormone
BMI	Body mass index
CI	Confidence interval
FSH	Follicle stimulating hormone
GNRH	Gonadotrophin-releasing hormone
HFEA	Human Fertilisation and Embryology Authority
ICC	Intraclass correlation
ICSI	Intra-cytoplasmic sperm injection
IUI	Intrauterine insemination
IVF	In-vitro fertilisation
LBR	Live birth rate
LH	Lutenising hormone
LLA	Lower limits of agreement
MoM	Multiples of means
NHS	National Health Service
NICE	National institute for health and clinical excellence

REC	Research ethics committee
RR	Relative risk
SD	Standard deviation
ULA	Upper limits of agreement
UK	United Kingdom

Chapter one: introduction to thesis

Chapter one: introduction to thesis

Desire to reproduce is a primal human instinct. Symbols demonstrating the cultural importance of human fertility can be seen throughout most ancient civilisations, showing that humans have been concerned about fertility from the onset of time (1). In the United Kingdom, one in seven couples experience an inability to conceive (2). Women facing fertility problems have reported higher levels of stress, anxiety and low mood symptoms (3). Globally, it is estimated that around 50-70 million couples experience difficulty with conceiving (4,5). This is likely to be an underestimate as there is poor access to fertility investigations and treatments in many developing countries (6).

Natural human reproduction requires the male to be able to produce sperm, achieve erection and ejaculation. It requires the couple to be able to have coitus. It requires the woman to be able to ovulate, to have patent and functioning fallopian tubes, to have a uterus which is structurally normal and an endometrium which is suitable for implantation (7). A problem with any one of these aspects could result in difficulties with conception. In about 25% of couples, despite being fully investigated for infertility, no cause is identified (8). This is referred to as unexplained infertility.

In vitro fertilisation

In-Vitro Fertilisation (IVF) is one of the treatment options available for couples struggling to conceive. The first baby conceived by IVF was in 1978 and Robert Edwards was awarded a Nobel prize in 2019 for this achievement (9). In the past 40 years, more than eight million babies have been conceived with the aid of IVF treatment worldwide (10). In 2014, in the UK alone, 52,288 women had IVF

treatment and 18,201 babies were born as a result (9). IVF is a suitable treatment option for couples with male factor infertility, damaged or blocked fallopian tubes, ovulation problems or unexplained infertility (11). IVF is a process by which a human egg is fertilised with a human sperm out of the woman's body. The fertilised egg is then placed inside the woman's uterus. Later in this chapter the IVF process is described in detail.

The cost of IVF

IVF treatment has cost implications. On average, an IVF cycle in the UK costs between £3000 to £8000 (12). It is estimated that half of IVF cycles are funded by the National Health Service (NHS) (12). Thus an unsuccessful IVF treatment has financial implications to the individual couple as well as the state. IVF treatment is also an intensive process requiring a woman to attend the hospital or clinic on multiple days. Therefore other costs associated with IVF, including the time off work taken by the couple and the financial implications to the couple, wider society and the state is difficult to estimate.

The cost of an unsuccessful IVF cycle cannot be quantified monetarily alone. Many studies have reported adverse mental health outcomes following unsuccessful IVF treatments (13–15). It has also been reported that the likelihood of a relationship break down is also significantly higher following unsuccessful IVF treatment (16).

The IVF process is also associated with some potential short and long term complications (17). As the number of cycles a woman undergoes increases, so does the likelihood of her encountering a clinical complication, such as acquiring an infection or damage to surrounding structures during egg retrieval.

Additionally, the chance of any particular IVF cycle resulting in a live birth is low. The work by McLernon et al., shows the probability of a live birth based on the age of the woman and the number of IVF cycles (shown in the Figure 1 below).

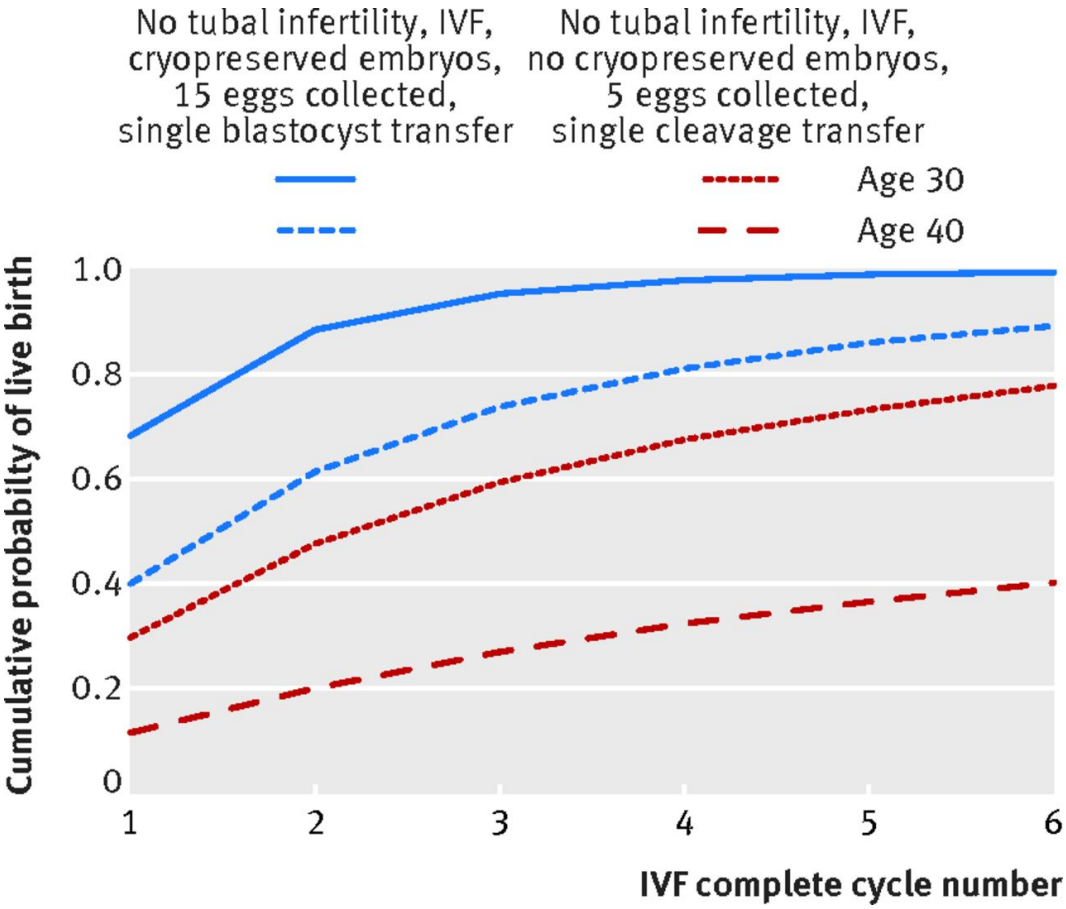


Figure 1: Probability of live birth per IVF cycle. (18)

It can be noted that a significant proportion of women will require multiple IVF cycles before achieving pregnancy. Some women may never achieve pregnancy despite multiple cycles of IVF treatments.

There is a clear benefit to the woman, the couple, society and the state in improving the success rate of each cycle.

The IVF process

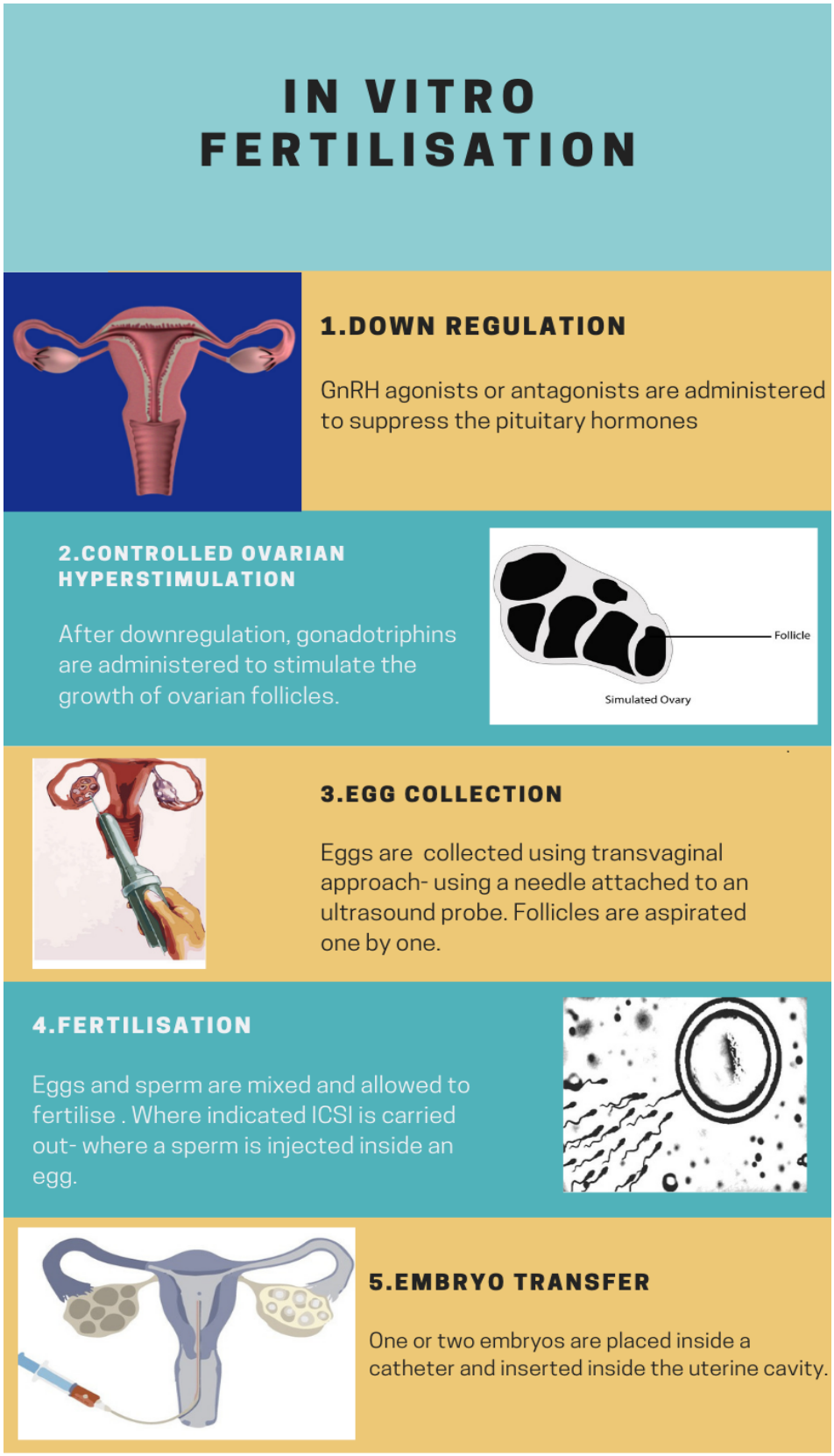


Figure 2: IVF key steps.

Significant disparity exists in the protocols used for IVF from one provider to another (19). The figure 2 above, summarises the key treatment steps in most IVF treatment regimes. In the following paragraphs I will explain the steps leading up to egg collection in more detail.

Pituitary down regulation

The existence of unpredictability, introduced by the possibility of ovulation before egg collection, could result in an inability to retrieve eggs and therefore unsuccessful cycles (20). To avoid the unpredictability of ovulation, as well as to enable scheduling, most clinicians carry out pituitary down regulation. Pituitary down regulation is where the production of luteinising hormone (LH) and follicle stimulating hormone (FSH) by the anterior pituitary gland is halted (21). This is achieved by either administering a gonadotrophin releasing hormone(GnRH)agonist, or a GnRH antagonist (21). A GnRH agonist acts by occupying all the GnRH receptors in the ovaries (22). A GnRH antagonist acts by blocking the GnRH receptors and thus blocking the release of gonadotrophins.

Controlled ovarian hyperstimulation

A fundamental objective of IVF treatment is to increase the yield of eggs produced. Multiple studies have demonstrated that the higher the number of eggs retrieved, the better the chances of achieving a live birth (23,24). This is because there is a gradual attrition of eggs or embryos along the IVF process. For instance, a significant number of eggs retrieved would not have reached sufficient maturity for fertilisation (25). Amongst the eggs that are mature, only just over half are likely to be successfully fertilised (26). Even after fertilisation, many fertilised eggs may not develop into embryos. From the embryos that are developed, their maturation may be halted before the stage at which embryos are considered ready for transfer into

the patient. Having a large number of eggs can therefore mitigate against this gradual attrition and thus improve IVF outcomes. If multiple embryos are available, then those with the best quality can be chosen for implantation. There is evidence that embryos that had been graded highly by an embryologist, had a higher chance of achieving a live birth (27).

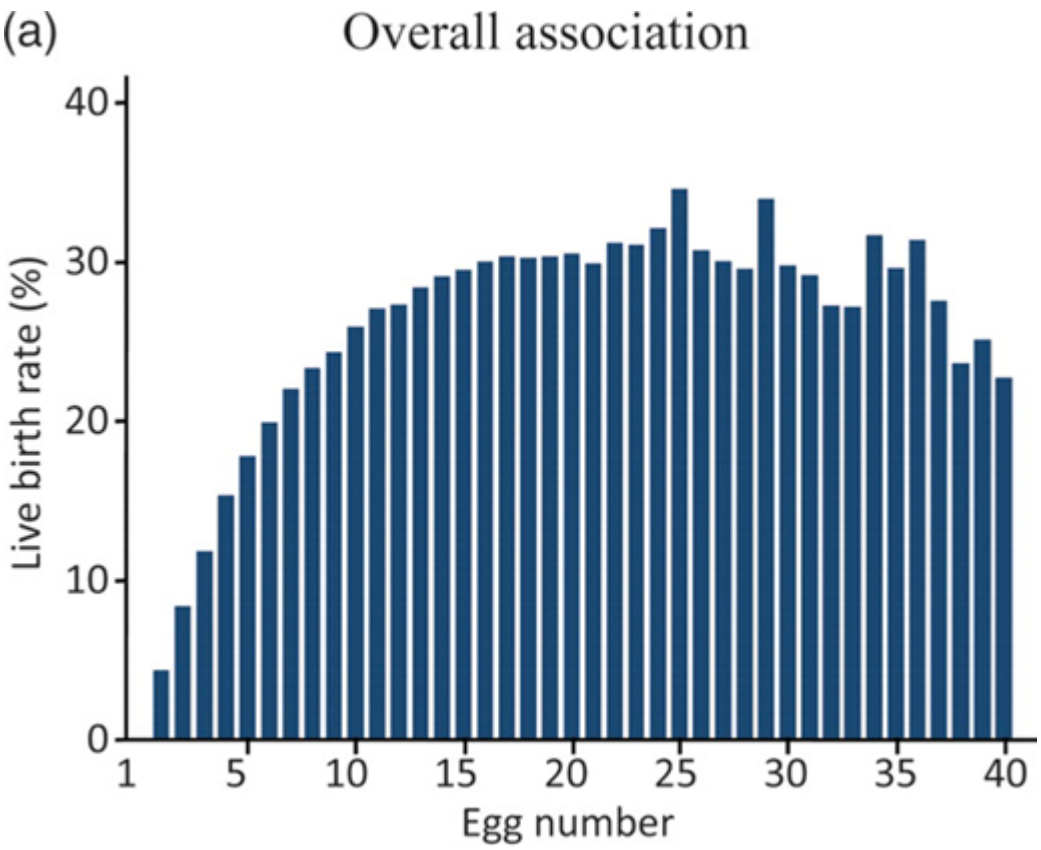


Figure 3- Association between egg number and live birth rate.(23)

The graph above (Figure 3), shows the work by Sunkara et al., demonstrating the relationship between the number of eggs extracted and the live birth rate. It can be seen that even a modest increase in the number of eggs extracted, for instance an increase from one to three, results in a higher than twofold increase in live birth

rate. Therefore increasing the number of eggs collected has the potential to improve reproductive outcomes significantly for women with low egg count.

Maximising the number of eggs collected is achieved by controlled ovarian hyperstimulation. This is the process initiated by the administration of gonadotrophins. Gonadotrophins are administered and follicular maturation is monitored by transvaginal ultrasound scans. Once a critical threshold is reached, which is usually having at least two follicles size above 15-18mm in size, an egg collection is planned (28).

Ovulation trigger

During physiological menstrual cycles, the rising level of oestrogens from the developing follicles results in the pituitary gland producing more LH. The LH surge results in the completion of the meiosis I in the oocytes and production of progesterone and prostaglandins within the follicle (29). The progesterone and the prostaglandins aid the breakdown of the follicle wall (30). This leads to the rupture of the follicle and release of egg. This event is known as ovulation.

Pituitary down regulation results in the absence of the physiological LH surge. Therefore the final egg maturation is achieved by an injection of human chorionic gonadotrophin (hCG), which mimics the action of the LH surge (31). Alternatively a gonadotrophin agonist injection (in those undergoing an antagonist protocol) can be administered, which leads to the production of the body's own LH surge (32). Subsequently eggs are harvested, inseminated and either embryos at cleavage stage or blastocysts are placed in the uterus.

The problem of poor response

As described above, a crucial step for IVF is controlled ovarian hyperstimulation in order to retrieve eggs. When an inadequate number of follicles develop in response to controlled ovarian hyperstimulation or when at the point of egg collection, an inadequate number of eggs are obtained, this leads to cancellation of that IVF cycle. This is known as poor response to controlled ovarian hyperstimulation. Some studies quote that up to 15% of women undergoing IVF do not respond to controlled ovarian hyperstimulation (33).

There are several risk factors associated with poor response, including increasing female age, history of previous poor response, low ovarian reserve and previous chemotherapy or radiotherapy (21). Low ovarian reserve is explained in subsequent paragraphs of this chapter.

Internationally there has been significant disparity in what is defined as poor response to controlled ovarian hyperstimulation. Because of the varying definitions, the exact number of IVF cycles that are cancelled due to poor response is difficult to state with accuracy.

The Bologna consensus statement by European Society of Human Reproduction and Embryology (ESHRE).

For a woman to be defined as poor responder to controlled ovarian hyperstimulation, she must meet two of the following criteria:

1. Advanced maternal age (defined as 40 years or more) or any other risk factor for poor ovarian response.

2. A previous cycle where there has been a poor ovarian response, defined as resulting in three or less oocytes with a conventional stimulation protocol

3.An abnormal ovarian reserve test (defined as an AFC of 5-7 or less, AMH of 0.5-1.1 pmol/L)(34).

Research also suggests that the percentage of women failing to respond to controlled ovarian hyperstimulation increases with advancing age (35). Given the fact that couples are choosing to delay conception and the mean age of patients undergoing IVF/ICSI is increasing, this problem is likely to increase in magnitude in the future (36). Hence there has been considerable interest in studying poor response, predicting poor response and studying the relationship between ovarian reserve and ovarian response (37,38).

Ovarian reserve

Ovarian reserve is the reproductive potential of a woman's ovary, thus refers to both the quantity and quality of the remaining oocytes (39). A woman has the highest number of oocytes as a fetus. This number steadily declines throughout life. Whilst age is a good predictor of the number of eggs, there can be significant difference in the number of eggs between two women of the same age (40). As a result, predictors

of ovarian reserve need to be utilised to estimate the ovarian reserve more accurately.

Ovarian reserve versus ovarian response

The terms ovarian reserve and ovarian response are often used interchangeably in literature, however they are different (41). Whilst ovarian reserve relates to the reproductive potential of a woman, ovarian response refers to how the ovaries respond to controlled ovarian hyperstimulation in a particular IVF cycle (42). It has been shown that there may be differing ovarian responses in the same woman who has multiple cycles of IVF treatment (43). Ovarian reserve tests are used by some to estimate the ovarian reserve as well as to predict the ovarian response (44).

There are a number of ovarian reserve tests used, which include measuring serum levels of FSH, oestradiol, inhibin B, anti-mullerian hormone (AMH), the clomiphene challenge test and AFC. Of these investigations, FSH, AFC and AMH are considered to be the most clinically useful investigations and are recommended for routine use by the NICE guidelines (8).

Follicle stimulating hormone

FSH is a glycoprotein polypeptide secreted by the anterior pituitary gland in response to gonadotrophin releasing hormone (GnRH), which is released by the hypothalamus. FSH production initiates follicle growth and maturation. Inhibin is produced by the granulosa cells and has a negative feedback on both the hypothalamus and the pituitary gland, inhibiting the production of FSH (45).

Serum FSH levels are measured during the early follicular phase of the menstrual cycle, namely days one to four. FSH is a relatively cheap and easy blood test to

perform. However, there can be significant intra and inter-cycle variation in FSH values (46). Whilst FSH is a reasonable predictor of ovarian reserve, some clinicians consider it an unreliable predictor of ovarian response to controlled ovarian hyperstimulation (47). The diagnostic sensitivity of measuring serum FSH to predict poor response varies from 10-80%, with a specificity quoted between 64-100% (39).

Anti-mullerian hormone

AMH is a glycoprotein secreted by primary, pre antral and antral follicles (48). AMH is produced by the ovarian granulosa cells. Studies have shown that AMH captures the decline in ovarian reserve better than FSH (49,50). There is also strong positive correlation between AMH and AFC (51). Studies have also shown that AMH can be a good marker for measuring ovarian responsiveness. AMH has been found to be lower in women who have had poor response to controlled ovarian hyperstimulation, in comparison to women with normal response (52).

Serum AMH has been shown to be relatively constant throughout the menstrual cycle. This gives AMH a presumed advantage over FSH and AFC, as timing measurement with the correct part of the menstrual cycle is not strictly necessary (53). Measuring serum levels of AMH is however expensive in comparison to FSH, and is consequently not available in most UK NHS hospitals.

Antral follicle count

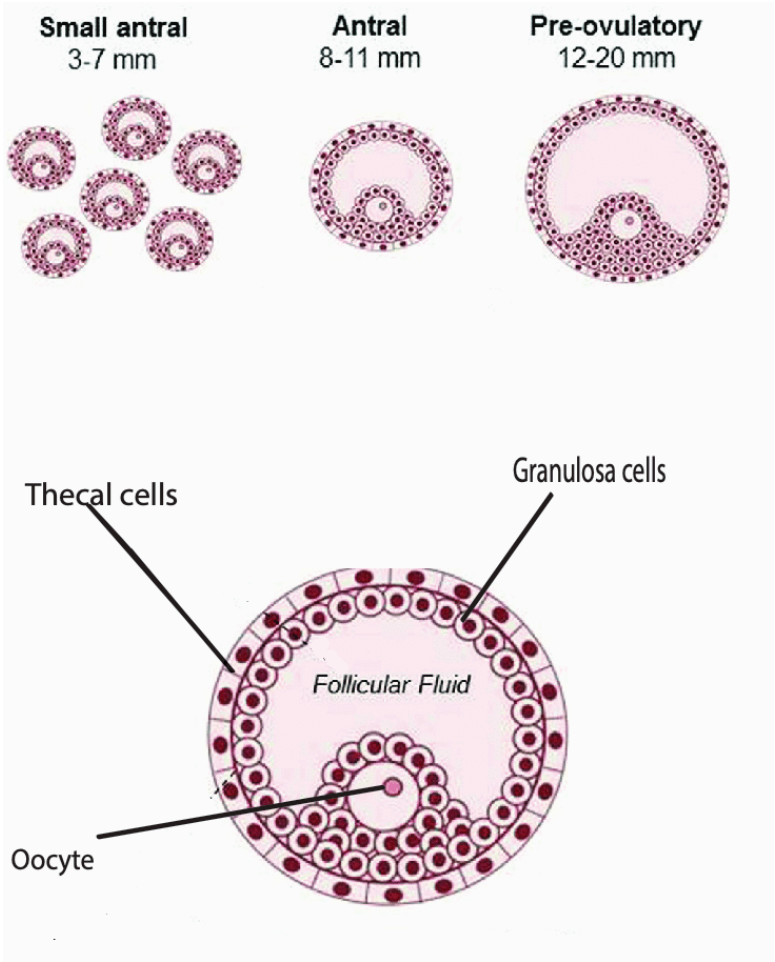


Figure 4: Antral follicle (Modified from Kristensen et al., with permission)(54)

The primordial follicle is an oocyte (egg cell), surrounded by a layer of oestrogen secreting cells called the granulosa cells. All primordial follicles are cells which have been paused in a stage of cell division called meiotic prophase (54).

At the beginning of each menstrual cycle, a number of primordial follicles get recruited to develop (55). The exact mechanism by which this occurs is poorly understood. It has been suggested that the number of follicles recruited correlates with the size of the total oocyte pool in the ovaries (56). It has been demonstrated that follicles are recruited in response to a rise in serum FSH levels (57).

The selected primordial follicles start to enlarge, and a layer called the zona pellucida, emerges. The granulosa cells in the follicle secrete oestrogen (58). At this stage the follicle is known as a pre antral follicle. The secreted oestrogen makes the wall of the egg cells increase the number of FSH receptors they have (59). The resulting increased FSH levels and the increased cell receptivity, lead to increased oestrogen production. The higher levels of oestrogens present, mature the pre-antral follicles to become antral follicles. Antral follicles measure between 2-10mm and can often be visualised by a transvaginal ultrasound scan (60).

AFC shows the follicles that are available for recruitment during controlled ovarian hyperstimulation. AFC is found to be a reliable reflection of the entire reserve of oocytes available (61). Studies have shown AFC to have a sensitivity of 9-73% and specificity of 73-100% for predicting ovarian response in an IVF cycle (39). Measuring and obtaining AFC is labour intensive, as it requires someone with expertise in ultrasonography. It also needs to be performed during the early follicular phase, making the timing of the test resource intensive. Some studies have also questioned the inter observer and intra observer variability of this test (62). Inter and intra observer variability of AFC will be explored further in chapter five.

Factors affecting ovarian reserve and reproductive outcomes

Advancing age is the biggest risk factor for diminished ovarian reserve and poor reproductive outcomes (63,64). This could be because of the decline of primordial follicles with age, which is described earlier in this chapter. It is also observed that with advancing of age, there is also a decline in the quality of oocytes(65). There have been limited human studies carried out studying the underlying mechanisms that result in poor quality of oocytes due to ethical and regulatory restrictions. Volarcik et al., studied donor oocytes and reported a greater proportion of meiotic

errors in oocytes obtained from older women(66). This study also found a greater degree of chromosomal segregation errors in oocytes obtained from women. Mice models have demonstrated an age related decline in the proteins cohesion and centromere specific histones, and as these proteins play an important role in chromosomal segregation, this may account for the greater degree of chromosomal segregation errors in older women(67).

Body mass index (BMI) is associated with ovarian reserve tests, with women with a higher BMI showing lower AMH levels, lower FSH levels and higher AFC (68–70). Lower FSH levels in obese women can be explained by the fact that adipose tissues produce oestrogens (70). Oestrogens have an inhibitory effect on FSH production through a negative feedback mechanism (71). Therefore women with high BMI who generally have more adipose cells are likely to have more circulating oestrogens, lowering the FSH levels. There are many mechanisms suggested for obesity's effect on AMH and AFC, including some studies linking obesity with reduction in granulosa cell activity and suppression of Inhibin B (70).

Other studies have suggested links between genes such as FMR1, ethnicity, smoking, vitamin D levels and caffeine intake on ovarian reserve tests (72–77).

Ovarian reserve markers and ovarian response

Ovarian reserve markers are considered to be modest predictors of ovarian response (44). Some studies have shown that AFC, AMH and FSH can be predictors of live birth (78–80). However the ability of ovarian reserve markers to predict live birth has been questioned by a more recent study (81). Studies have also shown that ovarian reserve markers are predictive of the number of eggs retrieved (82–84).

AFC in particular has been shown to have a linear relationship with the number of eggs retrieved (82,84).

Biological mechanisms

Numerous mechanisms have been postulated to explain the relationship between ovarian reserve tests and reproductive outcomes. Warburton et al had proposed the limited pool hypothesis(85). This theory suggests that when there is normal ovarian reserve, there are a number of follicles available at the beginning of each cycle to choose from. Therefore in natural conception the follicle containing the best quality egg becomes the dominant follicle by natural selection. In IVF, the ere are a number of embryos for the embryologists to choose from, and therefore the chances of a better quality embryo being chosen is higher. Conversely, in low ovarian reserve, there are fewer follicles and fewer eggs to compete to ovulate in natural conception, and there are fewer embryos to choose from in IVF. This could explain the lower rates of live birth and higher rates of pregnancy loss observed in women with low ovarian reserve. Ovarian reserve tests are also used by clinicians in determining the dose of FSH to administer to stimulate the ovaries, with women with low ovarian reserve getting higher doses of FSH(86). The work by Check et al., had demonstrated that higher doses of FSH adversely affects the quality of embryos in IVF cycles(87). This could also account for the relationship between ovarian reserve tests and reproductive outcomes.

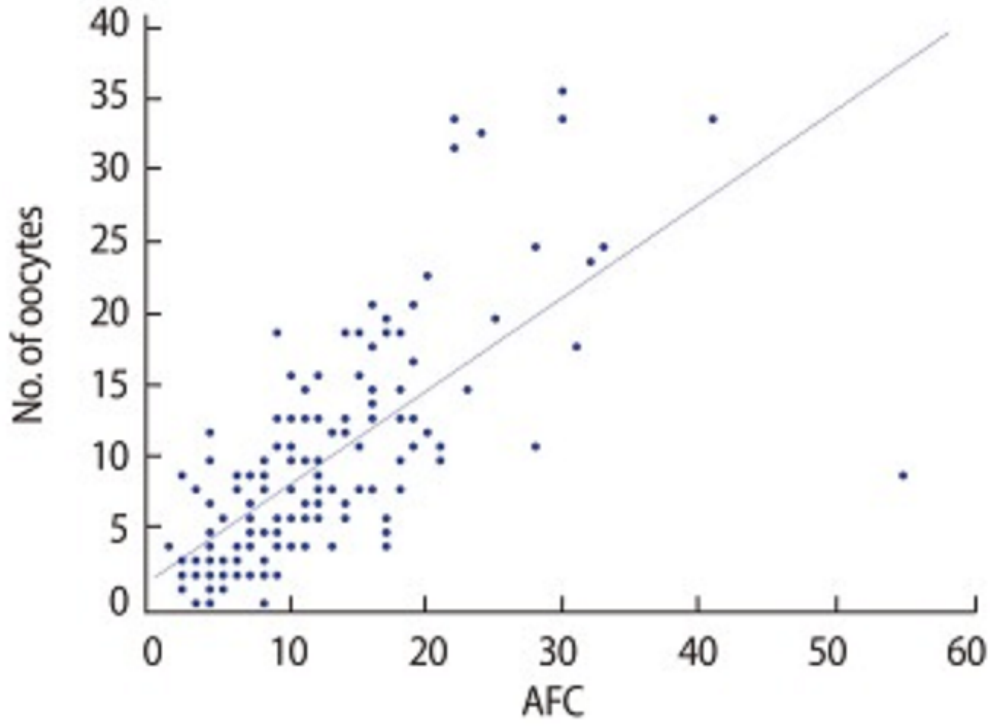


Figure 5 AFC and the number of eggs collected- Moon et al.(84)

The figure above is the work by Moon et al., demonstrating the linear relationship between AFC and the number of eggs collected (84). Even in women with a low AFC, it can be seen that a small increase in AFC leads to a corresponding increase in the number of eggs retrieved.

Inter-cycle variation of ovarian reserve markers

Inter-cycle variation refers to the difference in ovarian reserve test values in the same woman, during different menstrual cycles. Studies have suggested that there is inter-cycle variation in ovarian reserve test values in women (62)(88). Scott el al., postulated that this variation could be due to the fluctuation in circulating gonadatrophins (89). Animal studies have shown a direct link between administration of corticotrophin releasing hormone, which is secreted in response to stress, and an immediate decrease in the circulating pulsatile GnRH release (90). Other animal experiments have shown similar responses elicited by other stress

induced hormones (91). It can be postulated that the periodic variation of physical and emotional stressors experienced by women, may have an impact on their gonadotrophins, and ultimately their ovarian reserve tests.

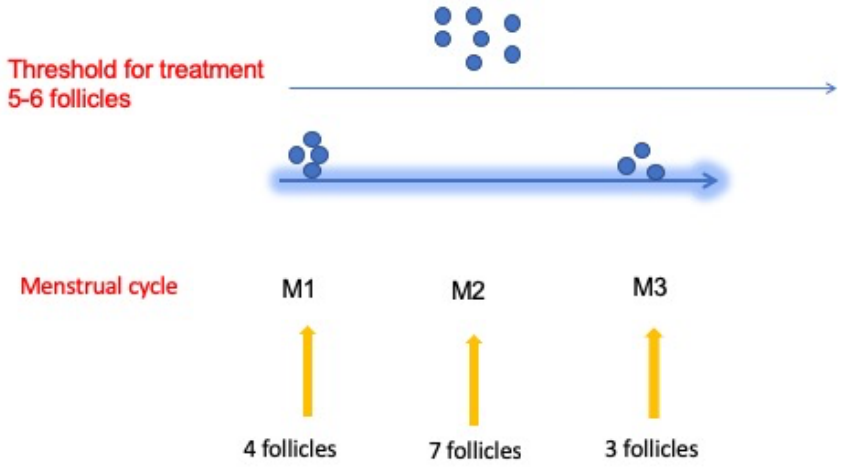


Figure 6: An illustrated example of cycle to cycle variation

I have explained in earlier paragraphs that there are fluctuations in ovarian response within the same woman. Some clinicians believe that the inter-cycle variation in ovarian reserve tests could be utilised to predict which monthly cycle was likely to result in the most favourable ovarian response. For example, in a particular month, there could be a low ovarian response, resulting in fewer eggs (for example four eggs), leading to fewer fertilised eggs (for example two eggs), further leading to only one embryo suitable for implantation. This could be contrasted with another month, in which the same patient has a better ovarian response, resulting in more eggs (for example seven eggs), leading to fewer fertilised eggs (for example four fertilised eggs) resulting in two or three embryos which are suitable for implantation. The question is firstly whether such variation in ovarian response exists from month to month, and secondly, how best to identify such variation.

Key messages from the introduction

Subfertility is a global problem which causes significant distress to those involved.

IVF is an effective option for many women with subfertility.

Ovarian reserve refers to the reproductive potential of a woman in terms of her egg reserve.

Ovarian response refers to how the ovaries respond to controlled ovarian hyperstimulation during IVF cycles.

Poor ovarian response is when the ovaries fail to respond to controlled ovarian hyperstimulation and do not produce sufficient eggs.

Poor ovarian response can have adverse consequences to the prospects of success with IVF treatments.

Ovarian reserve markers are tests used by clinicians to predict both the ovarian reserve and ovarian response.

It is likely there may be menstrual cycle to menstrual cycle fluctuation in ovarian reserve tests- known as inter-cycle variation.

It is the hypothesis of this thesis that inter-cycle variation in ovarian reserve tests could be used to identify the best month to start IVF treatment in women with low ovarian reserve, and potentially improve their IVF outcomes.

Aims and objectives of this thesis

In this thesis I set to understand if there is a menstrual cycle to menstrual cycle variation in ovarian reserve tests. The first step is to identify and collate the evidence that already exists. Currently there are no published systematic reviews on this topic. Therefore I carried out a systematic review on the intercycle variation in ovarian reserve tests, which I report in chapter two of this thesis.

My second objective was to identify the relationship between ovarian reserve tests and reproductive outcomes. As outlined earlier in this chapter, the relationship between ovarian reserve tests and live birth rate has been established in published literature. However there is no consensus on ovarian reserve tests and pregnancy loss, with no published systematic reviews on this subject. Therefore I carried out a systematic review and meta-analysis on ovarian reserve tests and pregnancy loss and recurrent pregnancy loss. I test the hypothesis that ovarian reserve tests correlate with pregnancy outcomes, independent of confounding factors such as age, ethnicity and BMI, by carrying out a primary study of retrospective analysis of IVF data.

My third objective was to identify if intercycle variation in ovarian reserve tests exist in women who have low ovarian reserve. I investigated this by carrying out a primary prospective cohort study in women with at least one risk factor for low ovarian reserve. I report these findings in chapter five of my thesis.

My fourth objective was to understand the patient perspective. I explored their views on having a low ovarian reserve in general, and in particular their acceptability of having a treatment based on antral follicle count and their

acceptability of taking part in a clinical trial to test the efficacy of such treatment. I report my thematic analysis of the prospective qualitative study I carried out in chapter six of my thesis.

My final objective was to gain insight into the clinician perspective of treating women with low ovarian reserve in general, as well as their their acceptability of a treatment protocol based on antral follicle count and their acceptability of testing the efficacy of such treatment protocol with a clinical trial. I report the thematic analysis of the qualitative study I carried out in chapter seven.

.

Chapter	Objective	Population studied	Design	Outcome of interest
2	To ascertain if menstrual cycle to menstrual cycle variation in AFC, AMH and FSH exists in women	Women of reproductive age	Systematic review	Inter-cycle variation
3	To ascertain the relationship between ovarian reserve tests (AFC, AMH and FSH) and pregnancy loss as reported in existing literature	Women having IVF treatment	Systematic review and meta-analysis	Pregnancy loss
4	To ascertain the relationship between AFC and reproductive outcomes	Women having IVF treatment	Cohort study	Live birth rate and pregnancy loss
5	To establish the inter-cycle variation in AFC, AMH and FSH in women with low ovarian reserve	Women with one or more risk factors for low ovarian reserve	Cohort study	Inter-cycle variation

6	To further understand the expectations of women with low ovarian reserve and to explore the acceptability of research and treatment based on inter-cycle variation	Women with a risk factor for low ovarian reserve	Qualitative study with purposive sampling and thematic analysis	Ideas, expectation and acceptability
7	To further understand the challenges faced by IVF clinicians in treating women with low ovarian reserve and explore the acceptability of research and treatment based on inter-cycle variation	IVF clinicians	Qualitative study with purposive sampling and thematic analysis	Ideas, expectation and acceptability

Table 1: Outline of thesis

Chapter two: inter-cycle variation of antral follicle count, anti-mullerian hormone and follicle stimulating hormone- a systematic review

Preamble to chapter two

In chapter one, I explained the burden of subfertility and the impact it has on women, couples and the wider society. I explained the problem of poor ovarian response and how ovarian reserve markers can predict poor response. I also referred to studies showing association between ovarian reserve tests, eggs retrieved and live birth rates in women who had assisted conception. I introduced the concept of menstrual cycle to menstrual cycle variability of ovarian reserve tests and the hypothesis of whether such inter-cycle variation could be used in the timing of IVF treatment and to improve outcomes for women with low ovarian reserve.

The first step in testing this hypothesis is to establish whether menstrual cycle to menstrual cycle variation of ovarian reserve tests exists in women.

Contributions

Dr Bala Karunakaran- Conceived the idea, carried out the search, collected the data, carried out quality assessment, analysed the data and wrote this manuscript.

Dr Abey Eapen was second reviewer for selecting included manuscripts and was the second quality assessor.

Mr Aurelio Tobias was consulted for his statistical expertise and advised that a meta-analysis was not possible.

Mr Jon Andrews was consulted for his expertise on search strategies.

Dr Ioannis Gallos and Prof Coomarasamy proof-read the manuscript and provided substantial edits.

Abstract

STUDY QUESTION

Is there variation in the ovarian reserve markers of AFC, AMH and FSH from one menstrual cycle to another, and if so what is the level of variation?

SUMMARY ANSWER

There appears to be variation from one menstrual cycle to another, with AMH showing the lowest amount of variation and AFC and FSH showing a higher degree of variation.

WHAT IS ALREADY KNOWN?

Multiple studies had commented on the inter-cycle variation of ovarian reserve markers, though very few studies had made direct comparison with each other. This is the first systematic review on this topic.

STUDY DESIGN, SIZE, AND DURATION

A systematic search of the literature was undertaken, according to the preferred reporting items for systematic reviews and meta-analysis (PRISMA) guidelines.

PARTICIPANTS/MATERIALS SETTING, METHOD

The following online databases were systematically searched: PubMed, EMBASE, CINAHL and Cochrane library up to 20 January 2018, with no language or date restrictions. The search terms included inter-cycle, inter-cycle, inter-cycle, menstrual cycle, variation, difference, and serial, follicle stimulating hormone, FSH, anti-mullerian hormone, AMH, antral follicle count, and AFC.

MAIN RESULTS AND THE ROLE OF CHANCE

A total of 253 studies were identified through the literature search. Eight studies met the inclusion criteria and data were extracted. A narrative synthesis was carried out. The trend was that AMH showed the least inter-cycle variability and FSH showed the most inter-cycle variability.

LIMITATIONS AND REASONS FOR CAUTION

Only one study had made direct comparison between AFC, AMH and FSH. Due to the heterogeneity of the way in which inter-cycle variation was reported, it was not possible to carry out a meta-analysis of pooled data. Therefore whilst trends can be observed, overall conclusions about inter-cycle variation cannot be made.

WIDER IMPLICATIONS OF THE FINDINGS

The existence of inter-cycle variation means clinicians should be cautious in forming judgements about a patient's ovarian reserve based on a single measurement of AFC, AMH or FSH. Further research is needed into inter-cycle variation in the poor responder and high responder cohorts and to see if this variation could be taken into consideration when planning individualised IVF protocols.

TRIAL REGISTRATION: Prospero CRD4201707442

Introduction

In 2015, in the United Kingdom, 72,504 In vitro fertilisation (IVF) cycles were carried out (92). One of the crucial steps in in vitro fertilisation (IVF) is controlled ovarian hyperstimulation (93). This is usually achieved by stimulation of the ovaries by gonadotrophins (94). The success of the stimulation, and the subsequent number and quality of oocytes retrieved plays a significant role in determining the success of the IVF cycle (23). Hence, significant research has been carried out in improving controlled ovarian hyperstimulation protocols to maximise the yield of oocytes retrieved.

Ovarian reserve tests such as follicle stimulating hormone (FSH), antral follicle count (AFC) and anti-mullerian hormone (AMH) tests have been proven to be surrogate markers for the success of IVF cycles (95). Ovarian reserve and ovarian reserve test values have been shown to decline with advancing age (96). The average age of a woman having IVF in the UK is 35, with the yearly trend showing an increase in the age at which women seek fertility treatment (92). Therefore there has been significant interest in optimising controlled ovarian hysterstimulation for women with low ovarian reserve. Many specialists also routinely use ovarian reserve tests to counsel women about their fertility potential, including commenting on their fertility potential outside the context of assisted reproduction (97). Therefore there is also a need to ascertain whether there is cyclical variability in ovarian reserve tests and whether a single result is sufficient for a clinician to confidently comment on a woman's ovarian reserve. Some studies have suggested that there is inter-cycle variation between the ovarian reserve markers (98). This review sets out to look at the inter-cycle variation of the three most clinically used ovarian reserve markers, AMH, FSH and AFC.

Materials and methods

The systematic review was conducted according to Preferred Reporting items for Systematic Review and Meta-Analysis (PRISMA) guidelines (99).

Inclusion and exclusion criteria

Criteria for inclusion were established prior to the literature search. The decision was made to include any cohort studies in which participants had more than one measurement of AFC, AMH or FSH, during the early follicular phase of their menstrual cycle. Early follicular phase was chosen as it is established in literature as the optimal timeframe to measure both AFC and FSH (100).Both prospective and retrospective studies were included and no language restrictions were place. Women were included irrespective of their fertility status.

Literature search strategy

A literature search was carried out on the following databases: MEDLINE, EMBASE, Cochrane central register of controlled trials and Web of science (inception - June 2017). A combination of Medical Subject Headings (MeSH) and text words were used to generate two subsets of citations. One subset included the following terms combined by the command OR: inter-cycle, inter-cycle, inter cycle, menstrual cycle, variation, difference, and serial. The second subset included the following terms combined by the command OR: follicle stimulating hormone, FSH, anti-mullerian hormone, AMH, antral follicle count, and AFC. The two subsets were combined by the command AND to generate a list of abstracts to be screened.

References of all included journal articles were examined to identify relevant articles not captured through the electronic searches. Primary authors were contacted for additional data and clarifications where possible.

Study eligibility and selection

All titles and abstracts were screened by two reviewers [Bala Karunakaran (BK), PhD student and Abey Eapen (AE) , PhD student] independently. A third reviewer [Ioannis Gallos (IG), subspecialist trainee in reproductive medicine] arbitrated if there were disagreements. Full manuscripts were obtained for all selected abstracts. Full manuscripts were reviewed by both reviewers independently first, and then together, and agreement on which studies to be excluded was reached. The Newcastle- Ottawa scale was used to assess the quality of the included articles. Newcastle-Ottawa is a numerical scale, with a maximum of nine stars awarded (101,102). The Newcastle-Ottawa scale awards four stars for study selection, two stars for comparability and three stars for study outcome. Each study received a score from each of the reviewers (see table below). There was provision for a third arbitrator (IG) in instances of disagreement, which was not necessary.

Data extraction and analysis

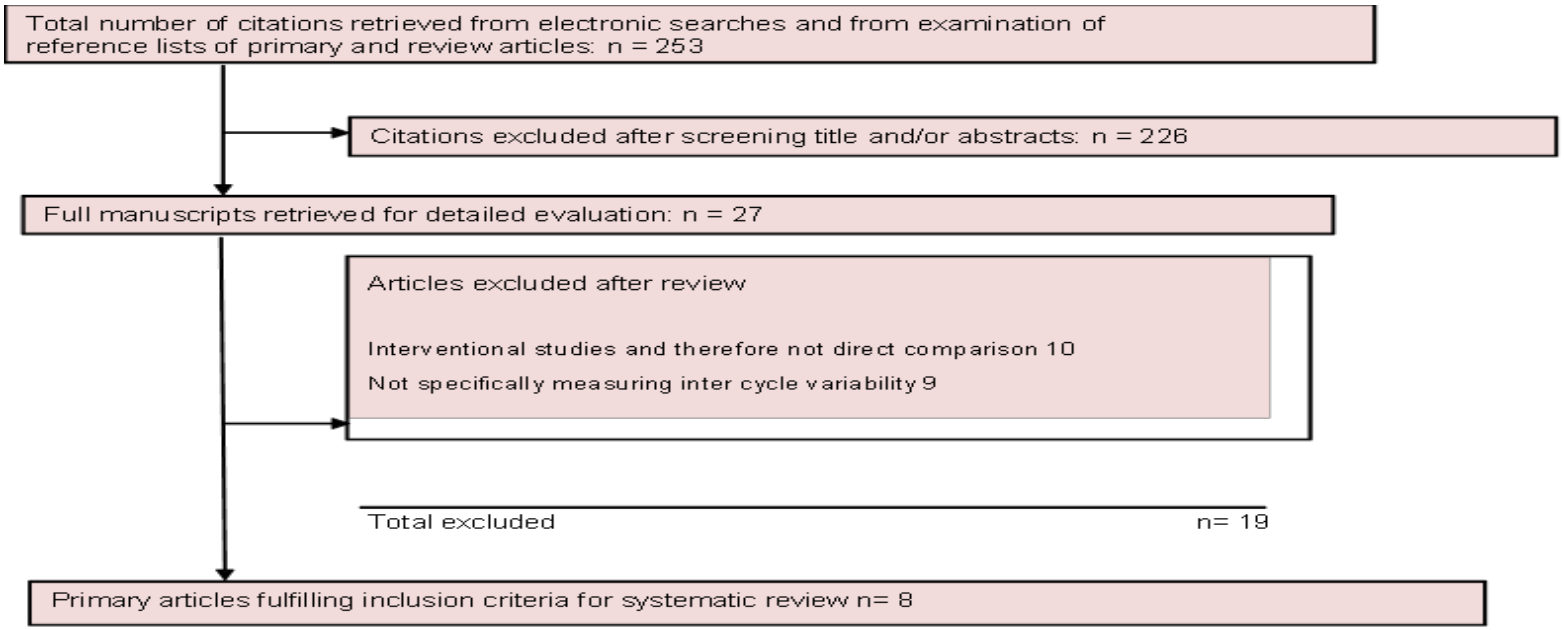
Data were extracted by the primary reviewer (BK) and tabulated. This was checked by a second reviewer (AE). The obtained information included study characteristics including - study design, study population, details of the tests, results and statistical analysis used. Data from the included articles of the systematic review are presented descriptively. As there was significant heterogeneity in the study populations, test processes, and test thresholds, a numerical pooling (meta-analysis) was not considered appropriate.

Results

Characteristics of included studies

Our literature search yielded 253 citations. 240 were identified through search of databases and 13 through screening of reference lists. The 253 titles and abstracts were screened by both reviewers (BK and AE). From that list 27 studies were chosen for full manuscript review. After a comprehensive review of manuscripts by both reviewers (BK and AE), eight studies were deemed to meet the criteria for inclusion in the systematic review. The study screening and selection is expressed in the form of a flow diagram (see Figure 6)

Figure 7 Study selection



Data extraction were performed on eight papers. Three of the studies assessed healthy women with no history of subfertility. The remaining five studies included women with subfertility. Seven of the studies were published in peer reviewed journals. One of the studies was a conference abstract (103) . Due to heterogeneity of the reporting of inter-cycle variation and lack of uniformity in the statistical expressions made, a cumulative synthesis and meta-analysis was not possible. Therefore a decision was made to carry out a descriptive analysis.

Quality of the included studies

The table below shows the quality of included studies. The studies varied in quality, ranging between five and six stars in the Newcastle-Ottawa scale. Most studies studied women with difficulties conceiving. Two of the studies, Bancsi et al., and Brown et al., studied women with no reported fertility issues. All of the studies, apart from Bancsi et al., were published in peer reviewed journals. The studies spanned over a period starting from the year 1992 (Brown et al.,) and 2011 (Rustamov et al.,).

Table 2: Quality of included studies

	Selection				Comparability	Outcome			Comments
Study author, year of publication	**Representative-ness of the exposed cohort**	**Selection of the non-exposed cohort**	**Ascertainment of the exposure**	**Demonstration that outcome of interest was not present at start of study**	**Comparability of cohorts on the basis of the design or analysis**	**Assessment of outcome**	**Was follow-up long enough for outcomes to occur?**	**Adequacy of follow up of cohorts**	**Total rating**
Bancsi et al., 2004	★	NA	★	N/A	★	★	★ ★	★ ★	6 ★ ★
Brown et al., 1992	Women without reported fertility problems	N/A	★	N/A	★	★	★	★	5
Elter et al., 2005	★	NA	★	N/A	★	★	★	★	6 ★
Escobar et al., 2010	★	N/A	★	N/A	★	Over a six year period. Age related decline could have significant bearing on results.	★	Retrospective analysis	4 ★

Fanchin et al., 2005	★	N/A	★	N/A	★	★	★	★		6 ★
Jayaprakasan et al., 2008	★	N/A	★	N/A	★	★	★	★	6	★
Rustamov et al., 2011	★	N/A	★	N/A	★	★	★	★		6 ★
Scheffer et al., 1999	Women without reported fertility problems	N/A	★	N/A	★	★	★	★	5	★

Inter-cycle variation results

Table 3: Inter-cycle variation results

Author	Population	Study design	Investigation	Results	Interpretation
Brown JR et al., (104)	48 healthy women with regular menstrual cycles,	Prospective cohort	FSH Blood samples obtained on day 3 over multiple non-consecutive cycles in one year period.	Mean coefficient of variation (CV) was 25.6. The standard deviation of CV was 21.4-29.9.	FSH showed a degree of inter-cycle variability,
Elter et al., (105)	52 healthy women with regular menstrual cycles and proven to be ovulatory with mid luteal phase progesterone.	Prospective cohort,	FSH, AFC taken in early follicular phase over two consecutive cycles,	FSH Mean difference 0.27(+-1.83). Lower limit of agreement -3.2 (CI -4.2 to -2.4). Upper limit of agreement 3.86 (CI 2.98 to 4.74). Multiples of Mean 0.99, AFC Mean difference -0.17 (=-3.41) Lower limit of agreement -6.85 (CI -8.45 to -5.24) Upper limit of agreement 6.50 (CI 4.90 to 8.10) Multiples of Mean 13.35,	Both AFC and FSH were shown to have variability. FSH had greater inter-cycle variability in comparison to AFC.

Escobar et al., (106)	36 women having fertility treatment.	Retrospective cohort over a 6 year period.	AMH taken during non-consecutive menstrual cycles.	12/36 patients had at least one AMH value diverging by >50%.	AMH showed degree pf variability.
Fanchin et al., (51)	46 women with regular menstrual cycles and a history of fertility problems.	Prospective cohort.	AMH, FSH, AFC taken over three consecutive menstrual cycles.	AMH Intraclass correlation 0.89 (CI 0.83 to 0.94) FSH Intraclass correlation0.55 (CI 0.39 to 0.71). AFC Intraclass correlation 0.73 (CI 0.62 to 0.84).	The greatest inter-cycle variation is seen in FSH. The least amount of inter-cycle variation is seen in AMH.
Jayaprakasan et al., (62)	88 women with regular menstrual cycle having IVF treatment.	Prospective cohort.	AFC and FSH taken during the cycle before IVF treatment. Non-consecutive cycles.	AFC Mean difference 0.16(SD1.98) Lower limit of agreement -3.71. Upper limit of agreement 4.03 Multiples of Mean 0.48. FSH Mean difference -0.08 Lower limit of agreement -4.52 Upper limit of agreement 4.36 Multiples of Mean 1.29.	FSH is shown to have a greater inter-cycle variation in comparison to AFC.
Rustamov et al., (103)	186 women with regular menstrual cycles and a history of fertility problems.	Retrospective cohort.	AMH and FSH taken during non-consecutive cycles.	AMH Coefficient of variation 28% (SD 3.6). FSH	AFC and FSH are shown to have a similar inter-cycle variation.

				Coefficient of variation 27% (SD2.0).	
Bancsi et al., (107)	120 women with regular menstrual cycle having IVF treatmen.t	Prospective cohort.	AFC On day 3 of two spontaneous menstrual cycles.	Mean difference 0.05.	A small inter-cycle variation is demonstrated. Smaller magnitude in comparison to other studies that studies AFC.
Scheffer et al., (108)	81 women with regular menstrual cycles and no history of fertility problems.	Prospective cohort.	AFC measured over three consecutive menstrual cycles.	Mean difference -0.28 Upper limits of agreement 8.37 (CI 7.41 to 9.33). Lower limits of agreement -8.93 (CI -7.41 to -9.89).	Inter-cycle variation in AFC is demonstrated.

Brown et al., (104) included 48 female volunteers with no history of fertility problems and regular menstrual cycles, under the age of 40. It was prospective and serum FSH samples were obtained over multiple cycles over a year. Brown et al., also obtained multiple samples of serum FSH in the same cycle, on days two, three and four. The samples were not necessarily obtained during consecutive menstrual cycles. The inter-cycle variability was reported in the form of coefficient of variation (CV). CV is a measure that compares the standard deviation of a value in comparison to the mean (109). The greater the CV, greater the variability that exists. Brown et al., reported a CV of 25.6% [95% confidence interval (CI) 21.4-29.9%] from one menstrual cycle to another. This could be contrasted with the CV of day to day variability within the same menstrual cycle, where the CV was 14.9% (CI 11.5%-18.3%). The existence of greater CV between menstrual cycles in comparison to within the same menstrual cycle is indicative of the existence of inter-cycle variation in serum FSH measurements.

Elter et al., evaluated 52 women with a history of subfertility with regular menstrual cycles. AFC and FSH were measured on day two of the menstrual cycle over two consecutive menstrual cycles (105). Inter-cycle variation was determined using a Bland Altman plot and by calculating limits of agreement (LOA). Bland Altman is a plot where difference between the two measurements is plotted against the average of the two measurements (110). The LOA is 1.96 times the SD above and below the mean of differences and 95% of observed values should fall within the two LOAs. The study standardised the values, by dividing the range by its mean. The value derived is known as multiples of mean (MoM) (111). Higher the MoM, greater the deviation from the mean and can be considered a surrogate marker for variation. Serum FSH had a MoM of 0.99 with an upper limit of agreement (ULA) of 3.86 and the lower limit of agreement (LLA) was -3.32. AFC had a MoM 1.45 and ULA of 6.50 and LLA of -6.85. The authors of the study did further analysis on the clinical significance of the inter-cycle variability of AFC. They defined an AFC of ten or above to correlate with a good ovarian response. The

cut-off of ten was determined by the authors based on the previous work by Fleming et al (112),. The study found that 58% of the subjects had less than ten follicles in one cycle and more than ten follicles in another cycle (33). The authors concluded that the inter-cycle variability of AFC was thus clinically significant. The authors did not carry out similar analysis on serum FSH values or commented on the clinical significance of inter-cycle variation of FSH.

Escobar et al., was a retrospective study (published as a conference abstract), which evaluated 36 women who had four or more fresh IVF cycles (106). AMH was measured at the beginning of each cycle. It was reported that at least a third of patients had an AMH value that varied by greater than 50%. However as these measurements were not consecutive and were measured over a six year period, the effect of decline of AMH over time could not be excluded. As this study is only a conference abstract, further information such as baseline characteristics was not available to critique.

Fanchin et al,. was a prospective study which evaluated 47 women with a history of subfertility with normal menstrual cycles (51). AFC, AMH and FSH were measured during the early follicular phase in three consecutive menstrual cycles. Inter-cycle variation was expressed in the form of Intraclass correlation (ICC). ICC calculates the consistency of conformity between two variables (113). Therefore smaller the magnitude of ICC, greater the inter-cycle variability. AFC had an ICC of 0.73 [confidence interval (CI) 0.62 to 0.84], AMH values exhibited an ICC of 0.89 (CI of 0.83 to 0.94) and FSH exhibited an ICC of 0.55 (CI of 0.39 to 0.71). These results indicate that FSH showed the greatest inter-cycle variability and AMH showed the least inter-cycle variability.

Jayaprakasan et al., prospectively studied 100 women with normal menstrual cycles, with a history of subfertility undertaking fertility treatment (62). Measurements of AFC and FSH were taken in the early follicular phase on the cycle preceding the IVF treatment cycle. The second

measurements were not taken on a consecutive menstrual cycle. However all measurements were taken in a period spanning a maximum of 12 months, thus limiting the impact of age related decline in ovarian reserve. The inter-cycle variability was reported in the form of LOA and MoMs. AFC results had a LLA of -3.71 and an ULA of 4.03 with a MoM of 0.48. FSH had a LLA of -4.52 and a ULA of 4.36 and a MoM of 1.29. These results indicate that FSH values have a greater inter-cycle variation in comparison to AFC.

Rustamov et al., retrospectively studied 186 women with a history of subfertility, who had two samples of AMH and FSH taken during non-consecutive menstrual cycles, within an upper limit of one year interval between them (103). AMH had a CV of 28% and FSH had a CV of 27%, showing consistent inter-cycle variation. However the study is limited by the fact that not all women included had normal menstrual cycles.

Bancsi et al, prospectively studied 120 women with a history of subfertility with regular menstrual cycles (107). AFC measurements were taken during the early follicular phase of two consecutive menstrual cycles. The difference between the two values were calculated and expressed in the form of mean difference. As the differences of AFC between any two measurements could be either positive or negative in magnitude, when all the differences are added, they are likely to cancel each other. Therefore the mean difference, which averages all the values, is likely to be low. Bancsi et al found the mean difference to be 0.05.

Scheffer et al, was a prospective cohort study which studied which looked at 81 women with regular menstrual cycles, no known fertility problems and history of proven fertility (demonstrated by a history of at least one pregnancy carried to term) (108). Early follicular phase AFC were measured during three consecutive menstrual cycles. The inter-cycle variation was reported as LOA. The mean difference between the measured AFC were -0.28. The LLA was -8.93 (95% confidence interval -7.41 to -9.33) and the ULA was 8.37 (95% confidence

interval was 7.41 to 9.33). This shows the existence of inter-cycle variation in AFC measurements, to a greater extent than expressed by both Elter et al and Jayaprakasan et al.

Discussion

In our systematic review we found a limited number of observational studies that found varying degrees of variation within and amongst the three tests of FSH, AMH and AFC. All three of these tests in the studies we found showed inter-cycle variation. The trend indicates FSH having the greatest inter-cycle variation and AMH having the least inter-cycle variation.

Fanchin et al, was the only study which made direct comparison between AFC, AMH and FSH (51). The results of Fanchin et al., indicate that FSH had the greatest inter-cycle variability and AMH had the least inter-cycle variability. Rustamov et al., made direct comparison between AMH and FSH, with both demonstrating similar inter-cycle variation, with AMH having a CV of 28% and FSH having a CV of 27%. Jayaprakasan et al., compared the inter-cycle variation of AFC and FSH. It was demonstrated that FSH values showed greater inter-cycle variation in comparison to AFC. Analysing the results of all of the available results together, a trend of FSH having the most inter-cycle variability and AMH having the least inter-cycle variability can be noted.

This study has many strengths. This systematic review is the first systematic review conducted, to our knowledge, examining inter-cycle variation of ovarian reserve markers. The strengths of our methodology include agreeing on the research strategy a priori and registering the protocol, using a robust search strategy including individually searching all relevant databases, and placing no language restrictions on the studies. All of the included studies were comprehensively assessed for quality using the Newcastle Ottawa scale. All apart from one study, Escobar et al, scored highly on the applicable domains of the scale.

The main weakness of this review comes from the degree of heterogeneity in which inter-cycle variation was reported: therefore pooling data and carrying out a meta-analysis was not possible. Another weakness of this review is the differences in study design of the included studies. Brown et al., and Scheffer et al., looked a population with no history of subfertility and therefore it is difficult to generalise their findings to women seeking fertility treatment. Four of the studies, Brown et al, Escobar et al, Jayaprakasan et al., and Rustamov et al., did not examine consecutive menstrual cycles. Escobar et al., had results spanning over a six year period. As ovarian reserve is known to decline with time, the impact of time elapsed between the measurements could not be eliminated (63,114). Escobar et al., and Jayaprakasan et al., had studied women who had pituitary down regulation and ovarian stimulation in between the measurements. There is limited published data available on the impact of ovarian reserve after IVF stimulation (115). The heterogeneity in the statistical tools used to express inter-cycle variation makes direct comparison between all studies not possible.

A weakness of the analysis is the fact that different studies are likely to have used different assays to measure AMH. Studies carried out measuring AMH values on the same sample but using different assays have produced different results(116,117). Craciunas et al., had suggested that any measurements of AMH done prior to the introduction of the Gen IIm assays in 2013 should be treated with caution(118). As all of the studies that reported on the intercycle variation of ovarian reserve markers were carried out prior to 2013, with varying sensitivity, specificity and comparability, our results should be treated with caution.

The biological plausibility of inter-cycle variation has been explored by studies. Scott el al suggested that this variation could be due to the fluctuation in circulating gonadotrophins, which have been shown in many studies (89). Animal studies have shown direct link between administration of corticotrophin releasing hormone , which is secreted in response to stress, and an immediate decrease in the circulating pulsatile GnRH release (90). Other animal

experiments have shown similar responses elicited by other stress induced hormones (91). It can be postulated that the periodic variation of stressors experienced by women, may have an impact on their gonadotrophins, and ultimately the function of the pituitary gland and the ovaries.

Only one study, Elter et al., explored the clinical significance of the inter-cycle variation and concluded that the inter-cycle variation in AFC was significant. This was achieved by broadly categorising all participants into two groups, those likely to have a normal response (AFC >/=10) and those likely to have a poor response (AFC < 10). Moving categories between menstrual cycles was deemed to be clinically significant. Caution is needed whilst interpreting this result, as a participant who had an AFC of ten in one menstrual cycle and an AFC of 11 in another, would have been classified as having a clinically significant result in this study. However in clinical practice, whether the addition of a single follicle would translate to an increase in live birth rate is questionable.

The existence of inter-cycle variation in ORTs would suggest that clinicians should take caution when counselling women about their reproductive health based on a single measurement. AMH may be the most reliable test to look at a snapshot of the ovarian reserve, as it consistently showed the least inter-cycle variation. Therefore if a clinician needs to base their assessment on one test alone, AMH is the most suitable candidate. However both FSH and AFC, which showed higher degree of inter-cycle variation, which may reflect the dynamic state of ovarian function in a particular cycle. This could mean that values of FSH and AFC may prove more useful when deciding whether or not to commence an IVF cycle at that point in time. FSH and AFC values measured in a specific cycle could also be utilised when deciding the dose of gonadotrophin.

The clinical impact of inter-cycle variation is likely to be in those at risk of having a poor response to controlled ovarian stimulation, such as women with low ovarian reserve. For this group, a difference in a few follicles may translate to a significant difference in the oocytes retrieved.

Further research, studying a population with risk factors for poor ovarian response, comparing the inter-cycle variation of AFC, AMH and FSH over multiple consecutive menstrual cycles is recommended. Research correlating ovarian reserve test measurements with IVF outcome would also be beneficial in determining the clinical significance of inter-cycle variation. Further research in identifying any underlying biological mechanisms which can explain the intercycle variation is needed. Understanding such mechanism might provide therapeutic options to enhance a woman's ovarian response and lead to better pregnancy outcomes.

Conclusion

There is evidence showing inter-cycle variation in FSH, AMH and AFC. There is insufficient evidence to determine which one of these ovarian reserve markers shows the greatest degree of inter-cycle variability. Available data indicates that AMH is likely to have the least amount of inter-cycle variability. However it must be noted that not all studies are in agreement with this. There is also limited evidence which suggests the existence of modest inter-cycle variation in both AFC and FSH. Further research is warranted to study ovarian reserve markers over multiple consecutive cycles for women with low ovarian reserve, to determine if significant inter-cycle variation exists.

Chapter three: ovarian reserve tests (FSH, AMH and AFC) and pregnancy loss: a systematic review and meta-analysis

Preamble to chapter three

In this thesis so far, I have explained the concepts of ovarian reserve and ovarian reserve tests. In chapter two of this thesis, I have shared the results of a systematic review I carried out demonstrating that inter-cycle variation exists in ovarian reserve tests within the same individual. In chapter one of this thesis I explore studies that had reported on the relationship between ovarian reserve tests and live birth in women having IVF. Pregnancy loss is an adverse outcome in women, with often significant consequences to the couple. The existing literature on ovarian reserve and pregnancy loss is conflicting. Therefore in this chapter I share the results of the first ever systematic review carried out exploring ovarian reserve tests and pregnancy loss.

Contributions

Dr Bala Karunakaran- conceived the idea, carried out the search, collected the data , carried out quality assessment, analysed the data and wrote this manuscript.

Mr Derrick Yates was consulted on search strategy.

Dr Rima Smith was second reviewer for selecting manuscripts and the second quality assessor. She also helped with checking the accuracy of the forest plots.

Prof Coomarasamy proof-read the manuscript and provided substantial edits.

Abstract

Study question

Is there any association between reduced ovarian reserve and pregnancy loss?

Summary answer

Reduced ovarian reserve measured by AFC, AMH and FSH is associated with increased risk of pregnancy loss. Insufficient data is available on the relationship between ovarian reserve and recurrent pregnancy loss.

What is known already?

Ovarian reserve has been shown to be a good prognostic marker for the success of assisted conception treatments. However, the evidence from studies exploring the relationship between ovarian reserve and the risk of pregnancy loss is conflicting.

Study size, design, and duration

A systematic review and meta-analysis investigating the relationship between ovarian reserve and pregnancy loss in women undergoing assisted conception treatment and women with a history of recurrent pregnancy loss.

Participants/materials, setting, methods

The review protocol was registered with PROSPERO (CRD42018099041). Two groups of women were included in this review; women undergoing assisted conception treatment and women with a history of recurrent pregnancy loss. Literature searches were conducted to retrieve studies in MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials and

CINAHL, for observational studies from inception until 15th of December 2018. The Newcastle-Ottawa scale for observational studies was used to assess the quality of included studies.

Main results and the role of chance

Twenty-one studies involving 27, 249 participants matched the inclusion criteria for women undergoing assisted conception treatment and were included in the meta-analysis. Data from eight studies reporting on the relationship between ovarian reserve and recurrent pregnancy loss were insufficient for meta-analysis. There was a significant association between reduced ovarian reserve and pregnancy loss for all three ovarian reserve markers: low AFC (RR 1.63 [95% CI 1.42-1.87]), low AMH (RR 1.29 [95% CI 1.02-1.84]), high FSH (RR 1.24 [95% CI 1.12-1.38]).

Limitations, reasons for caution

The present study evaluated couples undergoing assisted conception treatment. As such the results may not be applicable to women who conceive naturally. Data were insufficient for adjusting the analyses based on known confounding factors such as age and thrombophilia so their influence cannot be excluded.

Wider implications of the findings

Women with reduced ovarian reserve should be counselled about the higher risk of having a pregnancy loss following assisted conception treatment.

Introduction

Ovarian reserve tests are widely used by clinicians to assess women, counsel women about their reproductive health and plan in-vitro fertilisation (IVF) treatment (100). National Institute for health and Care Excellence (NICE) guidelines recommend clinicians using antral follicle count (AFC), anti-mullerian hormone (AMH) and follicle stimulating hormone (FSH) to assess women seeking fertility treatment, as these tests are considered to be the most clinically useful (8). There are several studies which show that ovarian reserve tests are useful in assessing the ovarian response, yield of oocytes during egg collection and live birth rate (44,119,120).

It is estimated that around 125, 000 women experience first trimester pregnancy loss in the United Kingdom, with over 40, 000 women requiring hospital admission (121). It has been shown that women who conceive with assisted conception have a higher incidence of pregnancy loss in comparison to women who conceive naturally (122,123). Pregnancy loss is associated with complications such as blood loss, sepsis, subfertility and sometimes results in maternal death (121). It has been reported that pregnancy loss can have lasting psychological consequences, affecting both the female and male partner (124,125). There is also evidence that women who conceive with assisted conception have a higher levels of psychological morbidity in comparison to women who conceive naturally (126). There is evidence that risk assessment and counselling prior to fertility treatment can help women who are more likely to experience adverse outcomes and can lead to a reduction in psychological symptoms (127).

Previous studies have been conflicting and there are no published systematic reviews that explore the relationship between ovarian reserve tests and pregnancy loss. Providing women with information regarding their chances of experiencing pregnancy loss will better inform their decision on whether to undergo assisted conception. Furthermore, it will enhance pre-treatment counselling by clinicians and also help to target further research to improve

outcomes. The aim of this review is study the relationship between AFC, AMH and FSH and pregnancy loss.

Methods

The protocol of our systematic review was published in Prospero (CRD42018099041) on sixth of June 2018 prior to commencing search.

Data sources

MEDLINE, (from inception to December 2018) EMBASE (from inception to December 2018) and CINAHL (from inception to December 2018) databases were searched electronically and Web of Science was used to search for grey literature.

The search of MEDLINE and EMBASE and CINAHL captured citations containing the relevant MeSH keywords and word variants were used to generate two subsets of citations. One subset was created combining the terms 'ORT', 'ovarian reserve*', 'AFC', 'antral follicle', 'follicle', 'FSH', 'follicle stimulating hormone', 'folitropin', 'AMH', 'anti-mullerian hormone' and 'anti mullerian hormone' using the command OR. A second subset was created combining the terms 'miscarriage', 'misc*', 'abortion', 'spontaneous abortion', 'abort*', 'pregnancy loss', 'live birth*' and 'pregnancy demise' using the command OR. The two subsets were combined with the command AND to generate a list of citation to be screened.

Bibliographies and reference lists of relevant articles were manually searched to identify papers not captured by the electronic searches. The 'similar articles' function in PubMed was used to identify further relevant publications. Authors were contacted where data were missing. There were no language restrictions placed.

Eligibility criteria for selecting studies

Primary observational studies which reported original data on association between ovarian reserve tests and pregnancy loss or recurrent pregnancy loss were included. Interventional studies, commentaries, narrative reviews, letters, case reports, conference abstracts and non-human studies were excluded. Studies were included regardless of year of publication, language or country of study.

Studies were selected in a two-stage process. Initially, all abstracts and titles were screened by two reviewers (Bala Karunakaran [BK] and Rima Smith [RS]) and full manuscripts of potentially eligible citations were obtained. A third reviewer (Arri Coomarasamy [AC]) was available to consult in instances of disagreement, which was not necessary.

Two subsets of studies were selected. One subset included studies where participants had a measurement of AFC, AMH or FSH and pregnancy loss rates were available. This included studies where pregnancy loss rates were reported, or where the reviewers were able to calculate the numbers of those who experienced pregnancy loss by subtracting the numbers of those who had a live birth from those with a positive pregnancy test. The second subset included studies examining any association between recurrent pregnancy loss and AFC, AMH and FSH.

Data extraction and synthesis

Data were extracted by two reviewers (BK and RS) independent of each other and then verified together. The primary outcome was pregnancy loss (as defined by the constituent studies).

Data were extracted from each paper for pregnancy loss rates and AFC, AMH or FSH categories. Where authors of the paper had classified a category as low ovarian reserve or normal ovarian reserve, we accepted their definitions. Where no such classification by the authors was available, we applied classifications of ovarian reserve as normal or low, trying to

closely match the reference ranges of other studies to make comparisons and meta-analysis possible.

The meta-analysis was performed using Review Manager (version 5.0 for Windows) to combine and analyse the data; using the generic inverse variance method. For the purpose of the meta-analysis, the ovarian reserve tests were grouped into three categories; AFC, AMH and FSH. In each category, the low ovarian reserve group was compared with the normal ovarian reserve group, against the outcome of pregnancy loss. Wherever possible, a subgroup analysis of the pregnancy loss rate in women with very low ovarian reserve was compared with women with normal ovarian reserve. Heterogeneity was assessed by examining the I^2 statistics. A random-effects model was used when there was statistically significant heterogeneity. Heterogeneity was presented statistically and graphically using forest plot estimates with 95% confidence intervals.

For studies investigating recurrent pregnancy losses, a narrative analysis was performed.

Quality assessment

All articles selected for meta-analysis were assessed for quality using the Newcastle Ottawa Scale (102). The Newcastle-Ottawa is a numerical scale, with a maximum of nine stars awarded (101,102). The Newcastle Ottawa scale awards four stars for study selection, two stars for comparability and three stars for study outcome. This tool was selected as it is designed to assess the quality of non-randomised studies; which is in-line with the study design of the included studies in this review.

Results

The search yielded 2,779 citations. Of these, 251 were duplicates and were excluded. 98 full manuscripts were obtained and screened. Of these, 68 manuscripts were excluded (pregnancy loss data not available [67 studies], donor egg recipients [1]). 21 studies were included in the meta-analysis and seven studies were included in the narrative review. The study selection process is summarised in

Figure **8**.

In total, 21 studies, consisting of 27, 249 participants, were included for the meta-analysis. Separate analyses were carried out for AFC, AMH and FSH. Two of the studies included for meta-analysis, Brodin et al., and Huyser et al., were prospective cohort studies (79,128). The rest of the 19 studies were retrospective cohort in nature. Tremellen et al., studied women having intra uterine insemination (IUI) treatment (129). The rest of the 20 studies were based on women who had IVF treatments. Included studies were conducted in nine different countries (United States [five studies], United Kingdom [four studies], Australia [four studies], Sweden [three studies], South Africa [one study], Taiwan [one study], Netherlands [one study], Brazil [one study] and Italy [one study]). The summary of study characteristics tables (Tables 8-10) below lists the definition of low ovarian reserve and normal ovarian reserve used in each paper. The quality assessment of included papers is summarised in Table **7**. The overall quality of included study was moderate.

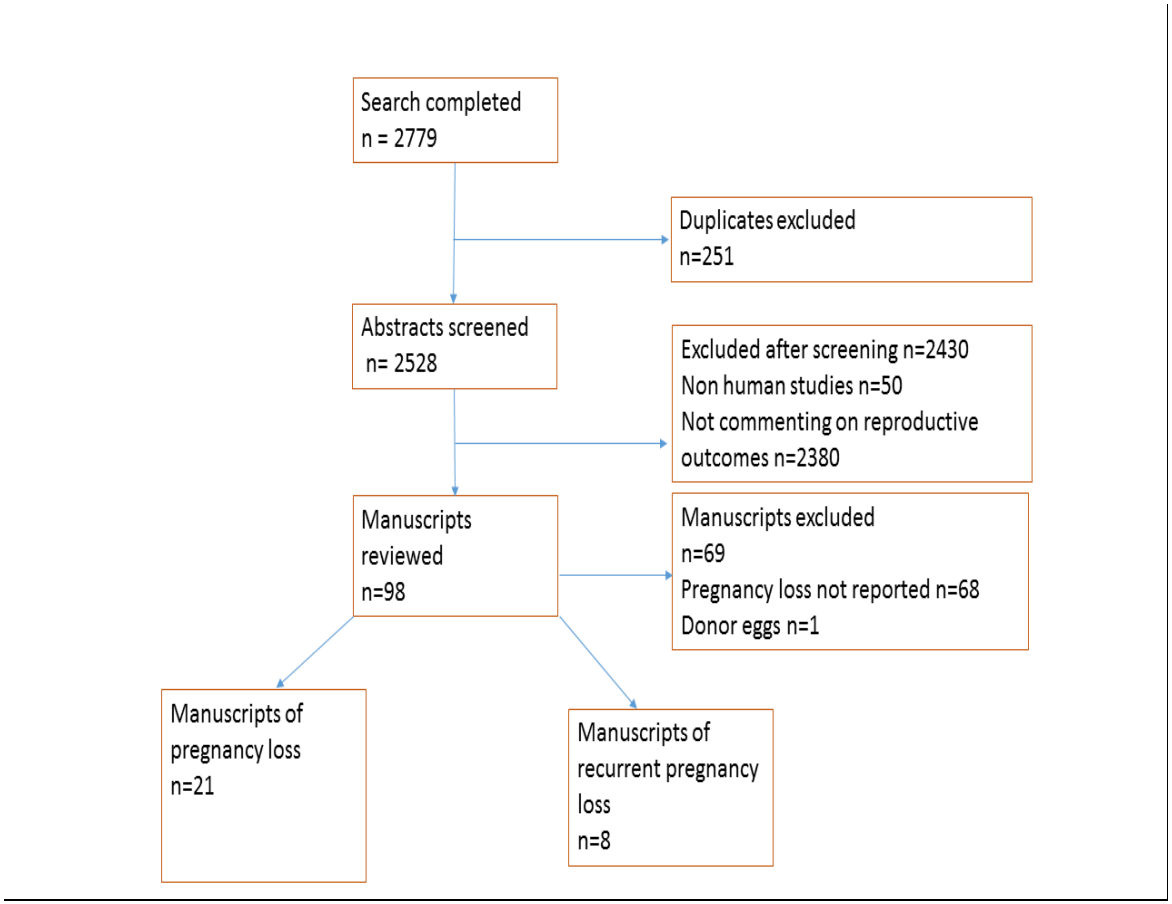


Figure 8: Selection process of included papers

Table 4: Characteristics of included FSH studies

Author (year)	Study design	Population	Definition of pregnancy loss	Definition of ovarian reserve categories	Number of women included
Abdalla et al., (2006)(130)	Retrospective cohort	IVF	Loss of pregnancy after a positive bHCG test and before completion of 24 weeks of pregnancy	Normal FSH ≤10IU/ml High FSH >10IU/ml	9
Bishop et al., (2017)(131)	Retrospective cohort	IVF	All women with a positive bHCG test who did not achieve live birth	Normal FSH ≤10IU/ml HighFSH >10IU/ml	8117
Caroppo et al., (2006)(132)	Retrospective cohort	IVF	All women with a positive bHCG test who did not achieve live birth	Normal FSH ≤10IU/ml High FSH >10IU/ml	19
Chuang et al., (2003)(133)	Retrospective cohort	IVF	All women with a positive pregnancy test who experienced a loss before 20 weeks gestation	Normal FSH ≤10IU/ml High FSH	395

				>10IU/ml	
Esposito et al., (2002)(47)	Retrospective cohort	IVF	All women with a positive pregnancy test who experienced a loss before 20 weeks gestation	Normal FSH ≤10IU/ml High FSH >10IU/ml	104
Huyser et al., (1995)(128)	Prospective cohort	IVF	Definition of pregnancy loss not reported	NormalFSH ≤11.68IU/ml High FSH >11.68IU/ml	40
Luna et al., (2007) (134)	Retrospective cohort	IVF	Pregnancy confirmed by ultrasound. No information on any gestational restrictions on the definition of loss provided	Normal FSH ≤13.03IU/ml High FSH >13.03IU/ml	2382
Sabatini et al., (2008)(135)	Retrospective cohort	IVF	Pregnancy confirmed by ultrasound and a loss before completion of 24 weeks.	Normal FSH ≤10IU/ml High FSH >10IU/ml	1589
Thum et al., (2009)(136)	Retrospective cohort	IVF	All pregnancies confirmed by ultrasound scan which did not result in live birth	Normal FSH ≤10IU/ml High FSH >10IU/ml	544
Toner et al., (1993)(137)	Retrospective cohort	IVF	Definition of pregnancy loss not reported	Normal FSH ≤10IU/ml High FSH >10IU/ml	48

Table 5: Characteristics of included AFC studies

Author (year)	Study design	Population	Definition of pregnancy loss	Definition of ovarian reserve categories	Number of women included
Holte et al., (2011)(78)	Prospective cohort	IVF	All pregnancies confirmed by ultrasound scan which did not result in live birth	AFC<11 low AFC>11 normal	804
Hsu et al., (2011)(138)	Retrospective cohort	IVF	All pregnancies confirmed by ultrasound scan which did not result in live birth	AFC<11 low AFC>11 normal	294
Keane et al., (2017)(139)	Retrospective cohort	IVF	All pregnancies confirmed by ultrasound but did not reach beyond 20 weeks gestation	AFC<9 low	423
Mustafa et al., (2017)(140)	Retrospective cohort	IVF	All pregnancies confirmed by bHCG but did not reach beyond 20 weeks gestation	AFC<9 low	534

Table 6: Characteristics of AMH studies

Author (year)	Study design	Population	Definition of pregnancy loss	Definition of ovarian reserve categories	Number of women included
Brodin et al., (2013)(79)	Prospective cohort	IVF	All pregnancies confirmed by ultrasound scan which did not result in live birth	Normal AMH>0.84ng/L Low AMH<0.83ng/L	337
Friden et al., (2011)(141)	Retrospective cohort	IVF	All women with a positive bHCG test who did not achieve live birth	Normal AMH≥8.6pmol/L Low AMH<8.6pmol/L	22
Keane et al., (2017)(139)	Retrospective cohort	IVF	All pregnancies confirmed by ultrasound but did not reach beyond 20 weeks gestation	Normal AMH≥8.6pmol/L Low AMH<8.6pmol/L	423
Lekamge et al., (2007)(142)	Retrospective cohort	IVF	All pregnancies confirmed by ultrasound. Any gestation limits for pregnancy loss not specified	Normal AMH≥14pmol/L Low AMH<14pmol/L	37
Pereira et al., (2016)(143)	Retrospective cohort	IVF	Pregnancy confirmed by ultrasound and a loss before completion of 24 weeks.	Normal AMH≥1ng/L Low AMH<1ng/L	500

Reijnders et al., (2016)(144)	Retrospective cohort	IVF	All women with a positive bHCG test who did not achieve live birth	Normal AMH≥1.05ng/L Low AMH<1.05ng/L	268
Tarasconi et al., (2017)(145)	Retrospective cohort	IVF	All pregnancies confirmed by ultrasound and did not carry on beyond 12 weeks gestation	Normal AMH≥1.60ng/L Low AMH<1.60ng/L	551≥
Tremellen et al., (2010)(129)	Retrospective cohort	IUI	All pregnancies confirmed by ultrasound and did not carry on beyond 12 weeks gestation	Age related percentiles to define low and normal	48

Table 7: Quality assessment of included studies (Newcastle-Ottawa Scale)

	Selection				Comparability	Outcome			Comments
Study Author, Year of Publication	Representativeness of the exposed cohort	Selection of the non-exposed cohort	Ascertainment of the exposure	Demonstration that outcome of interest was not present at start of study	Comparability of cohorts on the basis of the design or analysis	Assessment of outcome	Was follow-up long enough for outcomes to occur	Adequacy of follow up of cohorts	Total rating
Abdalla et al., (2006)	★	★	★	★	Not controlled for any confounders	★	★	Retrospective analysis	6 ★
Bishop et al., (2017)	★	★	★	★	Not controlled for any confounders. However age grouping available.	★	★	N/A as retrospective analysis	6 ★
Caroppo et al., (2006)	★	★	★	★	Not controlled for any confounders	★	★	N/A as retrospective analysis	6 ★

Chuang et al., (2003)	✹	✹	✹	✹	Not controlled for any confounders	✹	✹	N/A as retrospective analysis	6	✹
Esposito et al., (2002)	✹	✹	✹	✹	Not controlled for any confounders	✹	✹	N/A as retrospective analysis	6	✹
Huyser et al., (1995)	✹	✹	✹	✹	Not controlled for any confounders	✹	✹	N/A as retrospective analysis	6	✹
Luna et al., (2007)	✹	✹	✹	✹	Not controlled for any confounders	✹	✹	N/A as retrospective analysis	6	✹
Sabatini et al., (2008)	✹	✹	✹	✹	Not controlled for any confounders	✹	✹	N/A as retrospective analysis	6	✹
Thum et al., (2009)	✹	✹	✹	✹	Not controlled for any confounders	✹	✹	N/A as retrospective analysis	6	✹

Toner et al., (1993)	★	★	★	★	Not controlled for any confounders	★	★	N/A as retrospective analysis	6	★
Brodin et al., (2013)	★	★	★	★	Not controlled for any confounders	★	Loss to follow up not reported	★	6	★
Friden et al., (2011)	No description	★	★	★	★	★	Up to 20 weeks of gestation	★	6	★
Keane et al., (2017)	★	★	★	★	Not controlled for any confounders	★	★	N/A as retrospective analysis	6	★
Lekamge et al., (2007)	★	★	★	★	Not controlled for any confounders	★	★	N/A as retrospective analysis	6	★
Pereira et al., (2016)	★	★	★	★	Not controlled for any confounders	★	★	N/A as retrospective analysis	6	★

Reijnders et al., (2016)	✸	✸	✸	✸	Not controlled for any confounders	✸	✸	N/A as retrospective analysis	6	✸
Tarasconi et al., (2017)	✸	✸	✸	✸	Not controlled for any confounders	✸	✸	N/A as retrospective analysis	6	✸
Tremellen et al., (2010)	Somewhat representative as this studied patients having IUI	✸	✸	✸	Not controlled for any confounders	✸	✸	N/A as retrospective analysis	5	✸
Holte et al., (2011)	✸	✸	✸	✸	Not controlled for any confounders	✸	Loss to follow up not reported	✸	6	✸
Hsu et al., (2011)	✸	✸	✸	✸	Not controlled for any confounders	✸	✸	N/A as retrospective analysis	6	✸

Keane et al., (2017)	✹	✹	✹	✹	Not controlled for any confounders	✹	✹	N/A as retrospective analysis	6	✹
Mustafa et al., (2017)	✹	✹	✹	✹	Not controlled for any confounders	✹	✹	N/A as retrospective analysis	6	✹

AFC

The pooled meta-analysis for AFC is listed in Figure **9**. In total, there were five included studies with 3,413 participants. Results showed a statistically significant increased risk of having a pregnancy loss in women with low AFC compared to normal AFC; relative risk (RR) 1.63 (95% confidence interval [CI] 1.42-1.87). Heterogeneity was relatively low (I^2 = 16%).

A subgroup analysis of very low AFC, defined as studies that had AFC of five or less as a category (Holte et al. [AFC ≤5], Hsu et al. [AFC ≤4], Keane et al. [AFC≤4] and Mustafa et al. [AFC≤4]) was carried out (78,138–140). 694 women were included in the subgroup analysis with low heterogeneity (I^2 = 22%). Results showed a statistically significant increased risk of having a pregnancy loss in women with an AFC count of five or less compared to normal AFC; RR 1.62 (95% CI 1.23-2.14).

AMH

The results of the meta-analysis for AMH studies is listed in Figure **10**. There were eight included studies, with the largest being Tarasconi et al., with 1,153 participants (145). 3,098 women were included in the meta-analysis and there was moderate heterogeneity (I^2 = 44%). Results show a statistically significant increased risk of having a pregnancy loss in women with low AMH compared to normal AMH; RR 1.29 (95% CI 1.02-1.84). There were insufficient data to allow for a further subgroup analysis of very low AMH.

FSH

The results pf the meta-analysis for FSH studies is displayed in **Figure 14**. The largest study is Bishop et al., with 8,117 participants (131). In total, there were 11, 492 women in the ten included studies. There was very low heterogeneity (I^2 = 0%). Results show a statistically significant risk of having a pregnancy loss in women with high FSH compared to normal FSH; RR 1.24 (95% CI 1.12-1.38).

A subgroup analysis of very high FSH, defined as studies that had FSH of 14 IU/ml or more as a category (Bishop et al. (FSH ≥14 IU/ml), Caroppo et al. (FSH ≥ 15IU/ml) and Toner et al. (FSH ≥ 15IU/ml)) was carried out (131,132,137). In total, there were 7,264 participants in the three included studies. There was

very low heterogeneity (I^2 = 0%). Results show a statistically significant increased risk of having a pregnancy loss in women with FSH ≥ 14 IU/ml compared to normal FSH; RR 2.28 (95% CI 1.87-2.79).

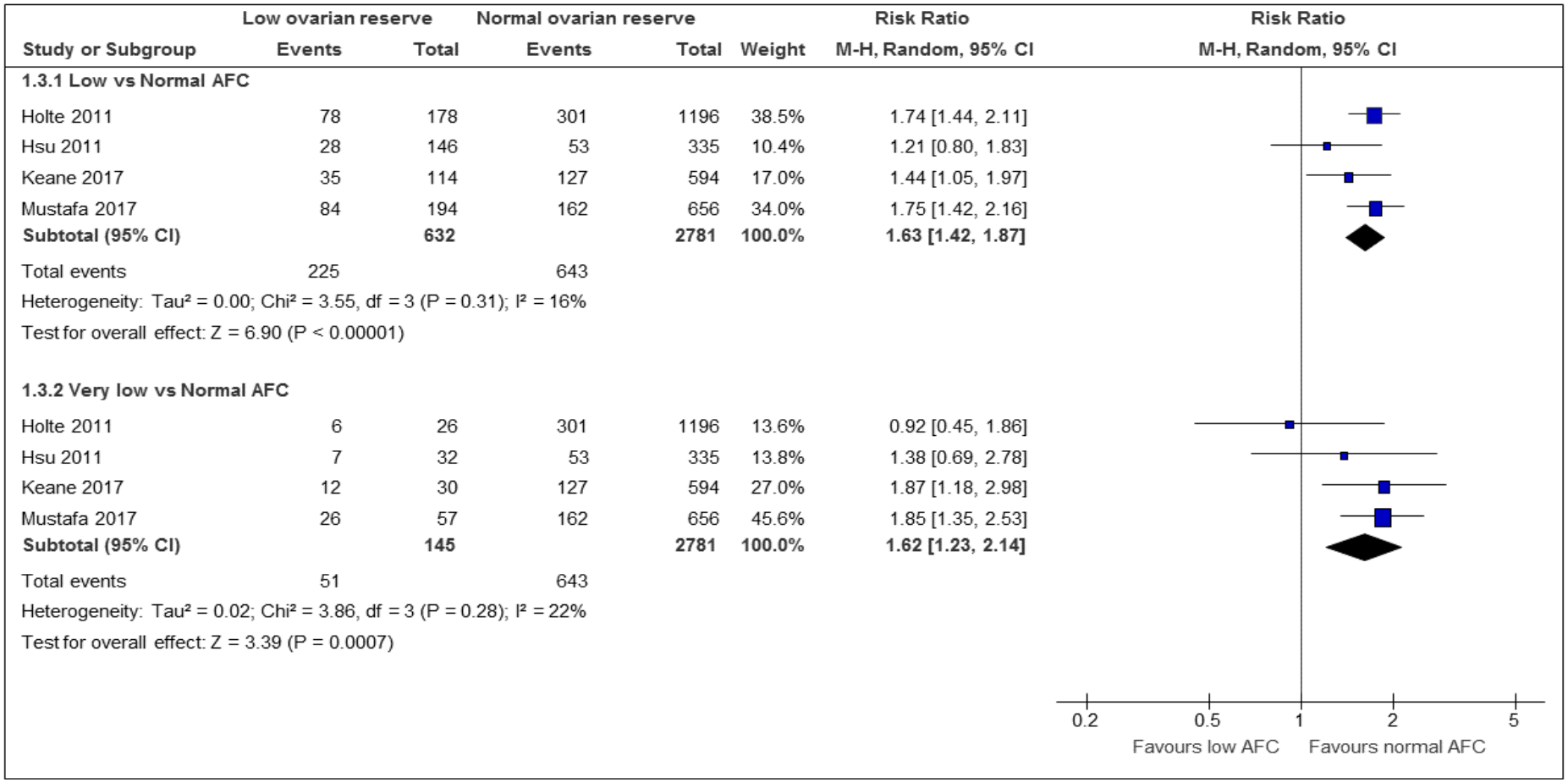


Figure 9: Forest plot showing pregnancy loss by AFC categories

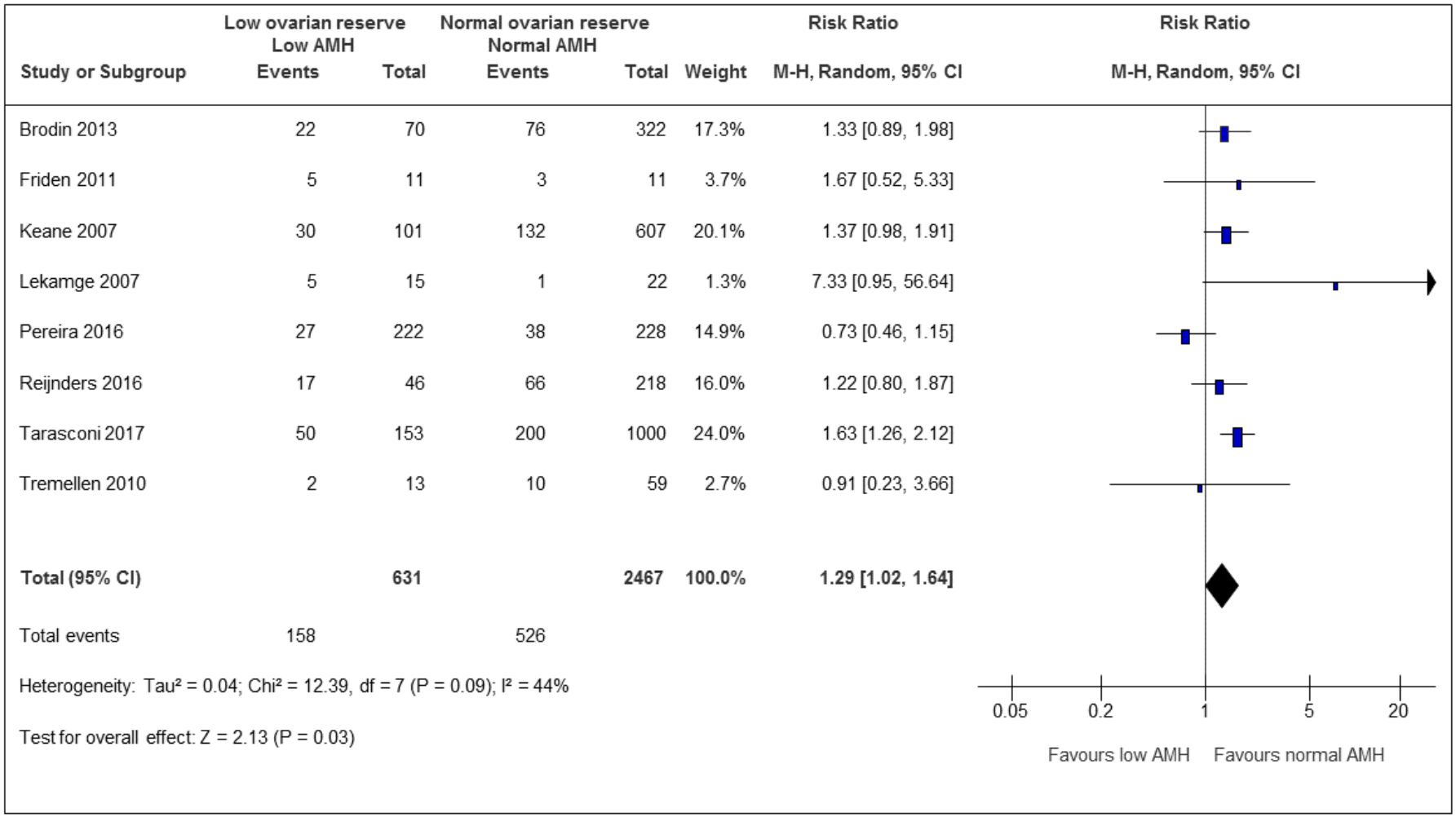


Figure 10: Forest plot showing pregnancy loss by AMH categories

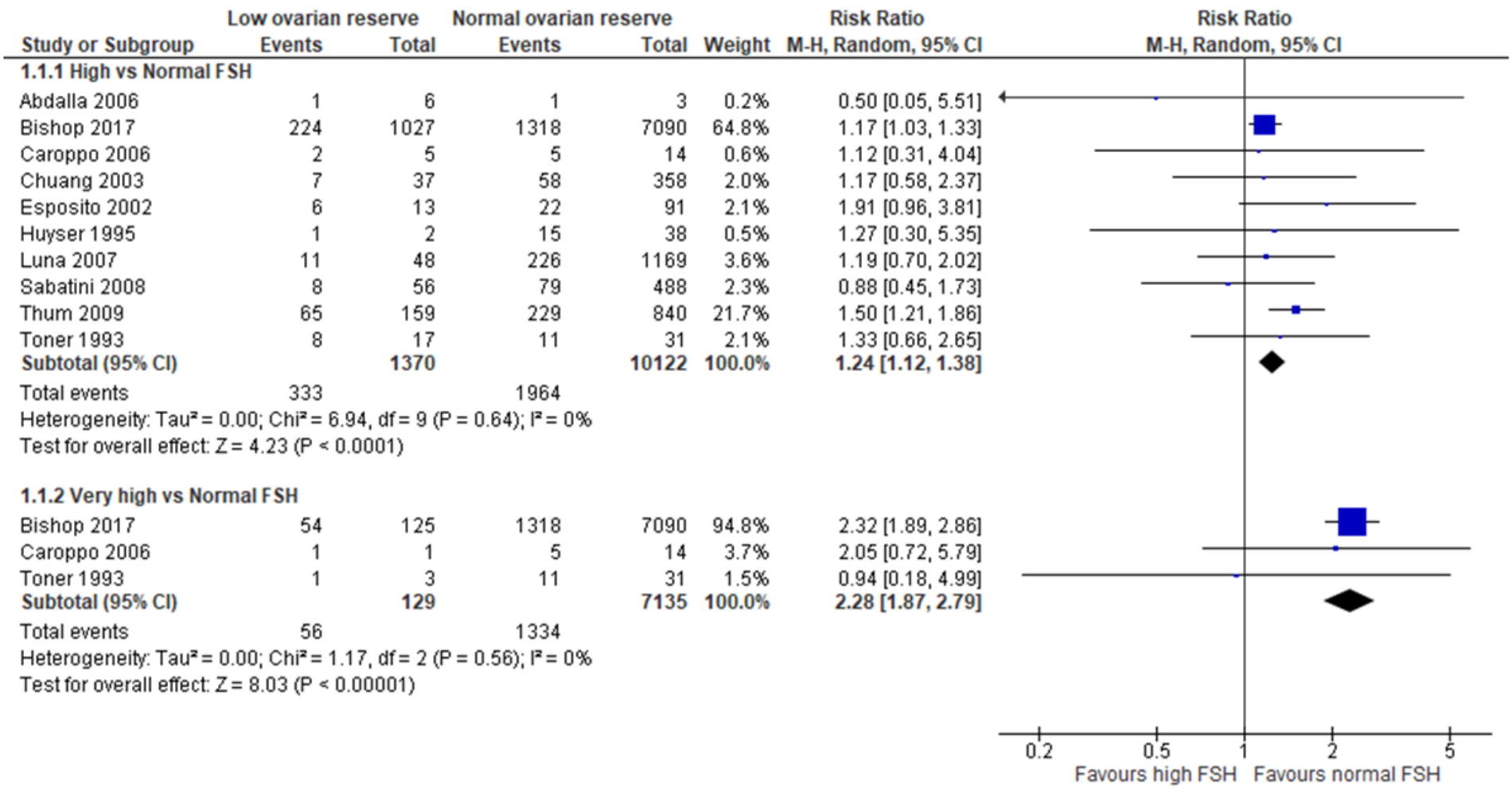


Figure 11: Forest plots showing pregnancy loss by AFC categories

Recurrent pregnancy loss

The studies that investigated women with recurrent pregnancy loss were heterogeneous. They varied in methodology and in how they reported their outcomes, and were unsuitable for meta-analysis.

Atasever et al., was a prospective cohort study carried out in Turkey over a five year period (146). The study compared women with a diagnosis of recurrent pregnancy loss, defined as three or more miscarriages, to a matched control group drawn from women seeking contraception. Atasever et al., measured AFC, AMH and FSH in the two groups and calculated mean values for each of the test and compared them. This study found that the mean levels of FSH were 8.6+-3.7UL (RPL group) and 7.1+-3.1U/L(control group), which was statistically significant (P=0.49). The levels of AMH were 2.9+-1.7ng/ml (RPL group) and 3.6+-1.6ng/L (control group), which was statistically significant (P=0.007). No significant differences were found in AFC measurements.

Bussen et al., was a case control study carried out at the USA, which compared women with a history of unexplained recurrent pregnancy loss, defined as three or more pregnancy losses, with a control group consisting of multiparous women attending a fertility clinic seeking treatment for either tubal infertility or male factor infertility (147). This study found no significant difference in the FSH levels.

Gurbuz et al., was a Turkish retrospective cohort study, which studied women with unexplained recurrent pregnancy loss (148). Recurrent pregnancy loss was defined as three or more consecutive losses. The control group of women included those with recurrent pregnancy losses, which were deemed as explained recurrent pregnancy losses by the investigators of the study. The mean FSH levels were 6.32

+-3.22 mIU/ml (explained RPL group) and 8.29+-2.99mIU/ml (unexplained RPL group which was statistically significant (P=0.007). However, the investigators found that the number of women with elevated FSH levels, defined as FSH>10mIU/ml, were evenly distributed between the two groups.

Hoffman et al., was a retrospective study based in the USA (149). It compared women with recurrent pregnancy loss and subfertility (pregnancy loss defined as three or more pregnancy losses), to women with subfertility who attended the same clinic. This study found no difference in the mean FSH concentrations of both cohorts.

Pils et al., was an Austrian retrospective cohort study (150). It studied women with recurrent pregnancy losses, defined as three or more consecutive pregnancy losses with the same partner. It compared women with unexplained recurrent pregnancy loss, to women with recurrent pregnancy loss and at least one causative factor as determined by the investigators. The study found the median AMH levels to be 2.0 ng/ml (explained RPL) and 1.2ng/ml (unexplained RPL), which was significant (p=0.037). No significant differences were found in FSH levels.

Prakash et al., was a study carried out in Sheffield, United Kingdom (151). It compared women with recurrent pregnancy losses, defined as three or more consecutive first trimester pregnancy losses, to women with no history of pregnancy loss or female factor subfertility. The study found no significant differences in mean AMH and FSH levels between the cohorts.

Trout et al., was a study carried out in the USA, studying women with recurrent pregnancy loss, defined as three or more first trimester pregnancy losses (152). Women with unexplained recurrent pregnancy los were compared to women with

recurrent pregnancy loss and at least one causative factor as determined by the investigators. The study compared the proportion of women who had elevated FSH levels, defined as ≥11mIU/mL. The percentage of women with elevated FSH levels was 31% (unexplained RPL) and 5% (explained RPL), which was significant (P=0.02).

Yuan et al., was a retrospective in design and it studied women in two different centres, one in China and one in the UK (153). It compared women with unexplained pregnancy loss with women who attended the fertility clinic in the same centres with no history of recurrent pregnancy loss. The investigators compared the proportion of women with high FSH levels, defined as ≥10IU/l, in the two cohorts. No significant differences were found.

The characteristics of the recurrent pregnancy loss studies are listed in Table **8**.

Table 8: Characteristics of recurrent pregnancy loss studies

Author (year)	Study design	Population	Ovarian reserve test studied
Atasever et al., (2016)	Prospective cohort	Women with recurrent miscarriage defined as three or more losses, compared with age matched controls.	AFC AMH FSH
Bussen et al., (1999)	Case-control study	Women with unexplained recurrent pregnancy loss defined as three or more losses. Comparison group consisted of multiparous women who were seeking fertility treatment for known tubal or male factor infertility.	FSH

Gurbuz et al., (2004)	Retrospective cohort	Women with unexplained recurrent pregnancy loss defined as three or more consecutive losses. The control group was women with recurrent pregnancy loss with what the investigators deemed a known causative factor, such as uterine septum	FSH
Hoffman et al., (2000)	Retrospective cohort	Women with recurrent pregnancy loss, defined as three or more losses, with subfertility, compared to women without recurrent pregnancy loss and subfertility.	FSH
Pils et al., (2016)	Retrospective cohort	Women with unknown recurrent pregnancy losses compared to with women with recurrent pregnancy loss with what was deemed by the investigators to be a causative factor. Recurrent pregnancy loss defined as three or more consecutive pregnancy losses, with the same partner.	AMH FSH
Prakash et al., (2006)	Prospective case control	Women with recurrent pregnancy losses, defined as three or more first trimester pregnancy losses. Compared to women with no history of pregnancy loss of female subfertility	AMH FSH

Trout et al., (2000)	Retrospective case control	Women with unexplained recurrent pregnancy loss compared to women with explained recurrent pregnancy loss. Recurrent pregnancy loss was defined as three or more first trimester pregnancy losses.	FSH
Yuan et al., (2012)	Retrospective cohort	Women with unexplained pregnancy losses, defined as three or more consecutive pregnancy losses, compared to women with a history of subfertility without recurrent pregnancy losses.	FSH

FSH and recurrent pregnancy loss

Five studies found no association between FSH levels and recurrent pregnancy loss. Prakash et al. found no difference between women who have had recurrent pregnancy loss and control group (151). Bussen et al., found no difference between women with recurrent pregnancy loss and the control group (147). Yuan et al. , Pils et al, and Hoffman et al. studied women with unexplained recurrent pregnancy loss and compared them with women with subfertility as controls, and found no association between FSH and recurrent pregnancy loss (149,150,153).

Three studies found an association between FSH levels and recurrent pregnancy loss. Atasever et al. compared women with recurrent pregnancy loss with matched controls and found that women with recurrent pregnancy loss have higher FSH levels (146). Trout et al. and Gurbuz et al. compared women with unexplained recurrent pregnancy loss with women with recurrent pregnancy loss with an identified causative factor, and found that women with unexplained recurrent pregnancy loss had higher levels of FSH (148,152).

AMH and recurrent pregnancy loss

Two studies found an association between recurrent pregnancy loss and low AMH. Pils et al. found that women with unexplained recurrent pregnancy loss had lower AMH levels on average in comparison to women with recurrent pregnancy loss with an identified causative factor (150). Atasever et al. found that women with recurrent pregnancy loss had lower AMH levels in comparison to matched controls (146).

We only found one study, Atasever et al., that explored recurrent pregnancy loss and AFC. No association was found.

Discussion

This was a comprehensive review of current evidence of association between the ovarian reserve tests of AFC, AMH and FSH and pregnancy loss. The meta-analysis of 21 included studies, suggests there is a positive association between reduced ovarian reserve, and increased risk of pregnancy loss, in women having assisted conception treatment. The risk of a pregnancy loss is increased in those with very high FSH in comparison to high FSH, indicating that a biological gradient of the effect may exist. No conclusions can be drawn about the relationship between AFC, AMH and FSH and recurrent pregnancy loss.

There are a number of possible theories to explain the relationship observed between reduced ovarian reserve and increased pregnancy loss. The limited pool hypothesis, suggested by many including Warbuton et al., explains that as those with low ovarian reserve are likely to have a lower number of follicles and yield less oocytes, the number of embryos available to pick for implantation is limited (85). Therefore, it is possible that clinicians are forced to use poorer quality embryos in those with low AFC.

Work by Grande et al. has shown a strong association between low AFC and rates of aneuploidy(129). It is possible that low ovarian reserve test values are surrogate markers of DNA damage of the oocytes. As the vast majority of early pregnancy losses are due to chromosomal errors, it is plausible that the higher incidence of pregnancy loss in those with low ovarian reserve could be attributed to the higher incidence of chromosomal errors (154).

Clinical implications

The findings of this study have clinical implications. When clinicians are counselling women with low ovarian reserve, they should not only counsel them about the anticipated lower IVF success chances, but also prepare them for the higher anticipated pregnancy loss rates. However, based on the current available evidence, we cannot recommend routinely testing for ovarian reserve in women experiencing recurrent pregnancy losses. We also cannot comment about the relationship between ovarian reserve tests and the chances of a natural conception resulting in pregnancy loss.

Strengths and limitations

This study has multiple strengths. To our knowledge this is the first systematic review on this subject. We followed a comprehensive search strategy, utilising multiple databases. The included studies scored moderately high on the Newcastle-Ottawa scale, with most of the studies losing stars in the Newcastle-Ottawa scale for comparability, as they failed to adjust for confounders and for follow up. Despite this assessment by the Newcastle-Ottawa tool, due to not adjusting the results for known confounders such as age, ethnicity and BMI, we deem the included studies to be of significant risk of bias.

Our study has many limitations. The main limitation is the heterogeneity among the studies in this review. The reference ranges applied to define low ovarian reserve were not consistent across the studies. Most of the studies were based on women with subfertility, thus limiting the generalisability of findings. None of the studies used standardised IVF protocols across all participants, thus making it difficult to rule out that differences in pregnancy loss, observed between the cohorts, cannot be accounted by treatment differences. Most studies did not measure the ovarian

reserve immediately before commencing the IVF protocol. As such, we cannot exclude the temporal effects on the recorded ovarian reserve test result and the actual ovarian reserve on the cycle of commencing treatment. The results are also not adjusted for known confounding factors, such as the woman's age and body-mass index. Not all studies used the same AMH assay. Liss et al., have demonstrated that the same serum sample could produce differing results based on the assay used (155). The impact of inter assay differences on AMH results cannot be excluded. Additionally, many of the studies had small sample sizes (Abdalla et al., n=9, Caroppo et al., n=19) (130,132).

Future research

A large prospective cohort study, exploring the relationship between ovarian reserve tests and pregnancy loss, with results adjusted for confounding factors such as age and BMI is recommended. Further research into whether donor oocytes, or starting the IVF treatment on a cycle with an optimum ovarian reserve test value may yield future treatment options for these women.

Chapter four: antral follicle count and the chance of livebirth and the risk of pregnancy loss in women having assisted conception: a cohort study

Preamble to chapter four

In this thesis I have explained what poor ovarian response is and the consequences it can have for affected women. I have explained the concepts of ovarian reserve and ovarian reserve tests. I have also shared studies showing that the higher the number of eggs retrieved, the better the chances of achieving a live birth. In chapter two of this thesis, I have shared the results of a systematic review I carried out, demonstrating that variation in ovarian reserve tests exists between one menstrual cycle to another.

As explained in chapter one, whilst there are some studies which show an association between ovarian reserve tests and live birth, there have been recent studies disputing the value of ovarian reserve tests in predicting that. Pregnancy loss is an adverse IVF outcome with sometimes devastating consequences. In chapter three I shared the results of a systematic review and meta-analysis I carried out- which showed an association between ovarian reserve tests and pregnancy loss. However the studies included in the systematic review had failed to adjust the data for known confounders.

In this chapter I describe a cohort study I carried out, studying the association between AFC and live birth rate and pregnancy loss, after adjusting for confounding factors such as age, ethnicity and BMI.

Contributions

Database was created and maintained by staff at CARE fertility.

Dr Bala Karunakaran- conceived the idea, the analysis strategy and wrote this manuscript.

Dr Ioannis Gallos helped with carrying out the analysis using Poisson regression.

Mr Aurelio Tobias helped with verifying the analysis using binomial regression and a generalised linear model.

Prof Coomarasamy and Dr Ioannis Gallos proof-read the manuscript and provided substantial edits.

Abstract

Study Question

What is the relationship between antral follicle count (AFC) and a) livebirth and b) pregnancy loss?

Summary Answer

Our study shows that AFC correlates with livebirth and pregnancy loss, independent of age, body mass index (BMI) and ethnicity.

What is known already?

Ovarian reserve tests (ORT), including AFC, have been shown to be good prognostic markers of the success of assisted conception treatments. However, studies evaluating the relationship between AFC and livebirth are limited, and studies exploring AFC and pregnancy loss have shown inconsistent findings.

Study size, design, and duration

The study included 25,767 cycles of in-vitro fertilisation (IVF) or intra-cytoplasmic sperm injections (ICSI). Data was collected prospectively from 2008 to 2016.

Participants, materials, setting, and methods

This was a study of women receiving IVF or ICSI treatment at any CARE (Centres for Assisted Reproduction and Embryology, UK) Fertility clinics in the UK and Ireland. Analysis was restricted to fresh embryo transfers following controlled ovarian stimulation. Frozen embryo transfers or transfers using donated oocytes were excluded. The primary outcomes were live birth rate (LBR) and pregnancy loss.

Main results and the role of chance

We included data from 25,767 cycles. AFC was available in 10,023 cycles. The live birth rate per embryo transfer was 33% (6,927/21,003). There was a significant increase in the live birth rate from 20% with an antral follicle count of 5 or less, rising to 41.1% with an antral follicle count of 25 or more (Figure 1).

Pregnancy loss rate, combining biochemical and clinical pregnancy losses, was 27.3% (2,656/9,719). AFC negatively correlated with pregnancy loss. Pregnancy loss was up to 41.7% with an antral follicle count of 5 or less and gradually decreased down to 23.2% with an antral follicle count of 25 or more. The association remained strong even when adjustments were made for age, body mass index and ethnicity for both live births and pregnancy loss.

Limitations, reasons for caution

This study evaluated couples with subfertility and studied women undergoing fertility treatment. Therefore, the results may not be applicable to women who conceive naturally.

Wider implications of the findings

AFC is a widely used test as part of investigating women who are struggling to conceive. This study shows that AFC can be an important aid for clinicians in counselling women about their likelihood of live birth and pregnancy loss and aid their decision making.

Introduction

The aim of fertility treatment is to enable the couple to conceive and raise a child. Therefore, live birth rate is considered to be the most important clinical outcome in assisted conception (156). Several studies have explored the relationship between AFC and livebirth in assisted conception. However, these studies were generally small, and were of variable methodological quality (78,157,158).

Pregnancy loss is a common complication, often with devastating consequences. Studies show that early pregnancy loss affects one in five pregnancies (159). Pregnancy loss is associated with significant psychological morbidity; it has been found that women may suffer with Post Traumatic Stress Disorder (PTSD) after experiencing the loss of their pregnancy (160), Amongst women who experience pregnancy loss, those who experienced a loss following IVF treatment are shown to have higher incidence of psychological symptoms in comparison to those who conceived naturally (126). Assisted reproduction is often self-funded by patients; as a result, there are often financial consequences following pregnancy loss post IVF treatment.

There is a clear need for clinicians to be able to counsel couples on their personalised chance of livebirth and risk of pregnancy loss before commencing IVF treatment. Being able to stratify the chance of livebirth and the risk of pregnancy loss according to AFC can ultimately help patients make informed choices about their treatment. There is evidence showing that in the IVF setting, when women with poor prognosis are adequately counselled and informed of the higher chances of adverse outcomes prior to commencing treatment, they coped better (161).

Antral follicles are the recruitable pool of follicles available in the follicular phase of the menstrual cycle. AFC is usually measured during the early follicular phase of the menstrual cycle (35).. AFC is commonly used as part of routine fertility investigations. Low AFC is established as a good predictor for poor response to ovarian stimulation in IVF (119).

There are conflicting results from studies which analysed AFC and live birth rate. Holt et al., demonstrated that AFC strongly correlated with live birth rate (78). Leijdekkers et al., demonstrated that

the addition of AFC to existing prediction models increased the ability for researchers to be able to successfully predict the chances of achieving live birth (158). However, the work by Hsu et al., showed that whilst AFC was useful in predicting the number of eggs retrieved in an IVF cycle, it was not useful in predicting live birth rates (162).

Few conflicting studies have been published exploring the relationship between AFC and pregnancy loss. Two cohort studies, Bishop et al., (131) and Haadsma et al., (163), found no association between AFC and first trimester pregnancy loss. However, Keane et al., another cohort study, reported an association between AFC and first trimester pregnancy loss. Thus, there is a clear need for further research to establish the relationship between AFC and pregnancy loss conclusively.

Material and methods

Study design

This study was an analysis of all cycles of IVF or ICSI treatment at any CARE fertility clinic in the UK and Ireland, from 2008 to 2016. CARE is a fertility provider which treats patients who self-fund their fertility treatment. The choice of treatment protocol and dose of medicines used at CARE is individualised to each woman and is left to clinicians' discretion. CARE database consisted of anonymised data extracted from prospectively collected clinical information as part of the routine care women received. Data were analysed from six fertility centres within the CARE consortium (Nottingham, Manchester, Northampton, Sheffield, Dublin and London) and a further seven satellite IVF units (Bolton, Boston, Derby, Leicester, Mansfield, Milton Keynes and Peterborough). This study protocol was developed based on the methodology described by Gallos et al., who had carried out analysis on a similar dataset (164). Analysis was restricted to fresh embryo transfer cycles following controlled ovarian stimulation. Frozen embryo transfer cycles were excluded from the analysis as there is a lack of consensus on the comparability of reproductive outcomes between fresh and frozen cycles. There is currently a large RCT being carried out to answer this question(165,166). Cycles using donated oocytes were also excluded. A separate sensitivity analysis based on women who had frozen embryo transfers or had oocyte donation could not be performed due to the small number of participants

Antral follicle count was measured during the early follicular phase (days 2-6) of the menstrual cycle before the IVF/ICSI treatment. The ultrasound scans were performed by sonographers, trained nurse sonographers or reproductive medicine specialists. IVF/ICSI treatment protocols were based on clinician discretion. Embryos transfers were performed by experienced clinicians using Wallace® Sure View® Catheter under ultrasound guidance.

Live birth was defined as the delivery of a live baby after 24 weeks gestation. As cycles that did not reach the stage of embryo transfer were excluded, the live birth rate was calculated per embryo transfer instead of per cycle started.

A pregnancy test was performed 18 days after oocyte recovery. Biochemical pregnancy loss was defined as a failure of pregnancy after a positive urinary pregnancy test, with or without visualisation of a pregnancy at first ultrasound examination (at 7 ± 1 week gestation) (167). Clinical pregnancy loss was defined as pregnancy loss after detection of fetal heart activity before 24 weeks of gestation. The outcome of pregnancy loss includes the combination of biochemical and clinical pregnancy losses.

Statistical analysis

Baseline patient characteristics, cycle characteristics and outcome data were described as frequencies with percentages, or means with standard deviations, as appropriate. The rates of the reproductive outcomes were plotted graphically using mean proportions and 95% confidence intervals.

A multiple-imputation analysis was adopted for any missing data for AFC, or covariates such as age and body mass index (29). The multiple imputation method imputed 10 values to fill in each of the missing values for antral follicle count or covariates. After that, statistical analysis was performed on the 10 imputed datasets separately and the results combined. The goal was to obtain better estimates of parameters and their standard errors (29). Dependant variable/ outcomes of livebirth follow a binomial distribution. However, as the sample size is large, it is possible to use Poisson regression as an approximation for binomial distribution. Advantage of such approach is that Poisson regression can be used to express the results as relative risk(RR), in oppose to odds ratio (OR). RR is a preferred measure

by many clinicians and is found to be more intuitive(168). A Poisson regression model was fitted to estimate crude and adjusted risk ratios for confounding variables such as age, body mass index and ethnicity. The assumption that the results produced by Poisson regression were valid was verified by comparing OR for all the values and RR for unadjusted early and late pregnancy loss, which were calculated using a binomial regression and a generalised linear model. This is included in Appendix 5.

The work by Dhillon et al., had shown that IVF outcomes could be confounded by factors such as age, ethnicity, BMI, parity and duration of infertility(169). In the subset of women with AFC results available, there were very few women with parity and duration of fertility accurately recorded. Therefore we adjusted the data for age, ethnicity and BMI which were reliably recorded in most women. We carried out a sensitivity analysis taking into account the within patient variability for women with more than 1 IVF/ICSI cycle in the database and adjusted for clustering. Further sensitivity analyses were performed to evaluate the impact of imputed values. All statistical analyses were done using Stata statistical software, release 14.

Ethical approval

Permission for use of the database was granted by CARE. Analysis of the database was anonymised. Such analysis of existing routine data supplied under license/agreement involving interventions in use only do not require formal IRB approval. However, each clinic is licensed by the Human Fertilisation and Embryology Authority (HFEA) and all activity is regularly inspected. The dataset was anonymized according to the ICO's (Information Commissioner's Office) guide on non-identifiable data. Furthermore, the CARE data protection certificate allows for the data to be used for survey and research purposes.

Results

There were 45,279 cycles performed at the CARE centres during the study period. We included data from 25,767 cycles after excluding cycles using donated oocytes, frozen embryo cycles, and cycles that were cancelled prior to embryo transfer. We chose to present the outcomes per embryo transfer rather than per cycle initiated, as the database did not contain sufficient information on the clinical indication for cancelling cycles. AFC was available in 10,023 cycles (46.1%) and these were included in the

regression model. All cycles were followed up to the point of the clinical pregnancy scan. The live birth outcome was not available for 752 cycles (3.5%) and they were excluded from the analysis for this outcome. The baseline characteristics are displayed in the table below. The mean AFC was 17.9 (Standard Deviation (SD) 12.8).

Table 9: Baseline characteristics

Characteristics	n (%) or mean (SD) Whole dataset (**n = 21,755)**
Age	34.8 ± 4.7
Duration of subfertility (years)	3 ± 2.2
Body Mass Index	24.7 ± 4
Ethnic background	
White	16,478 (83.1)
Asian	1,886 (9.5)
Black	242 (1.2)
Chinese	164 (0.8)
Other/Mixed	1,060 (5.4)
Antral follicle count	17.9 ± 12.8
Previous live birth	4,336 (19.9)
Previous miscarriage	4,790 (22)
Oocytes retrieved	9.6 ± 4.9
Ovarian stimulation protocols	
Long agonist	11,704 (54.1)
Short flare agonist	1,459 (6.7)
Antagonist	8,484 (39.2)
Cause of infertility	
Unexplained	5,173 (23.8)
Tubal	2,004 (9.2)
Ovulatory	1,139 (5.2)
Male	6,734 (30)
Uterine/peritoneal	935 (4.3)
Mixed or other cause	5,770 (26.5)
Treatment type	
IVF	5,920 (27.2)
ICSI	15,835 (72.8)

Table 10: Baseline characteristics of women with available AFC data

Characteristics	n (%) or mean (SD) Women with AFC data available (n = 10, 023)
Age	34.8 ± 4.6
Duration of subfertility (years)	2.7 ± 2.0
Body Mass Index	24.5 ± 3.9
Ethnic background	
- White	8,133 (83.2)
- Asian	877 (8.9)
- Black	138 (1.4)
- Chinese	82 (0.8)
- Other/Mixed	542 (5.5)
Antral follicle count	17.9 ± 12.8
Previous live birth	1,945 (19.4)
Previous miscarriage	2,465 (24.6)
Oocytes retrieved	9.5± 4.7
Ovarian stimulation protocols	
- Long agonist	6,859 (68.9)
- Short flare agonist	517 (5.19)
- Antagonist	2,584 (25.9)
Cause of infertility	
- Unexplained	1,796 (17.9)
- Tubal	857 (8.5)
- Ovulatory	432 (4.3)
- Male	3,371 (33.6)
- Uterine/peritoneal	391 (3.9)
- Mixed or other cause	3,176 (33.7)
Treatment type	
- IVF	2,823 (28.2)
- ICSI	7,200 (71.8)

AFC was split into categories in increment of five follicles to see if a gradual increase in AFC will result in a gradual increase in live birth rate and a gradual decrease in pregnancy loss rate. The live birth rate per embryo transfer was 33%. A decision was made to not analyse AFC based on clinical thresholds of low, normal and high as there is no clear clinical consensus on the cut offs for these labels. There was a significant increase in the live birth rate from 20% with an antral follicle count of 5 or less, rising to 41.1% with an AFC of 25 or more (see Table 11).

Table 11: Live birth and pregnancy loss per AFC category

Threshold	Live birth rate (95% CI)	Unadjusted RR (95% CI)	Adjusted RR (95% CI)	Pregnancy loss rate (95% CI)	Unadjusted RR (95% CI)	Adjusted RR (95% CI)
≤5	20.0 (16.92 to 23.43)	0.52(0.44 to 0.62)	0.47 (0.37 to 0.60)	41.7 (35.0 to 48.7)	1.85(1.57to 2.27)	1.90(1.50 to 2.39)
6-10	25.0 (23.00 to 27.0))	0.65 (0.59 to 0.72)	0.75 (0.67to 0.84)	32.6 (29.1to 36.2)	1.44 (1.22 to 1.70)	1.32(1.09 to 1.59)
11-15	31.2 (29.23 to 33.12)	0.81(0.74 to 0.89)	0.86 (0.78 to 0.96)	28.9 (26.0 to 31.8)	1.23(1.09 to 1.50)	1.22(1.02 to 1.47)
16-20	38.4 (36.01 to 40.90)	Reference	Reference	22.6 (19.7 to 25.6)	Reference	Reference
21-25	37.45(34.64 to 40.31)	0.97(0.88 to 1.07(	0.94(0.94 to 1.01)	26.9(23.4-30.7)	1.19 (0.99 to 1.43)	1.27(1.03 to 1.56)
>/25	41.12 (39.03 to 43.23)	1.07 (0.99 to 1.15)	1.04 (0.92to 1.10	23.2 (20.8-25.7)	1.03 (0.87 to 1.21)	1.11 (0.92 to 1.34)

Table 12: Early and late pregnancy loss per AFC category

Threshold	Early pregnancy loss rate (95% CI)	Unadjusted RR (95% CI)	Adjusted RR (95% CI)	Late pregnancy loss	Unadjusted late pregnancy	Adjusted late pregnancy
≤5	29.38 (23.33 to 36.03)	1.75(1.35 to 2.27)	1.84 (1.34 to 2.52)	12.32 (1.35 to 3.37)	2.13(1.57to 2.27)	2.02 (1.23 to 3.33)
6-10	19.39 (16.51 to 22.54))	1.15 (0.93 to 1.43)	1.14 (0.89to 1.46)	13.17 (10.74 to 15.92)	2.28 (1.62 to 3.20)	1.72(1.19 to 2.50)
11-15	19.33 (16.92 to 21.93)	1.15(0.94 to 1.40)	1.15 (0.91 to 1.45)	9.57 (7.81 to 11.57)	1.65(1.18 to 2.33)	1.39(0.96 to 2.02)
16-20	16.81 (14.28 to 19.59)	Reference	Reference	9.57(7.81-11.57)	Reference	Reference
21-25	19.36(16.25 to 22.77)	1.15(0.92 to 1.44)	1.25(0.97 to 1.63)	5.77 (4.26-7.62)	1.31 (0.88 to 1.95)	1.29(0.84 to 1.99)
>/25	15.72 (13.69 to 17.92)	0.94 (0.76 to 1.15)	1.05 (0.83to 1.33)	7.44(6.01 t 9.08)	1.28 (0.91 to 1.82)	1.24 (0.85to 1.82)

The overall pregnancy loss rate, combining biochemical and clinical pregnancy losses, was 27.3% (3308 /10 073). Pregnancy loss was up to 41.7% with an AFC of 5 or less and gradually decreased down to 23.2% with an antral follicle count of 25 or more (see Table 11).

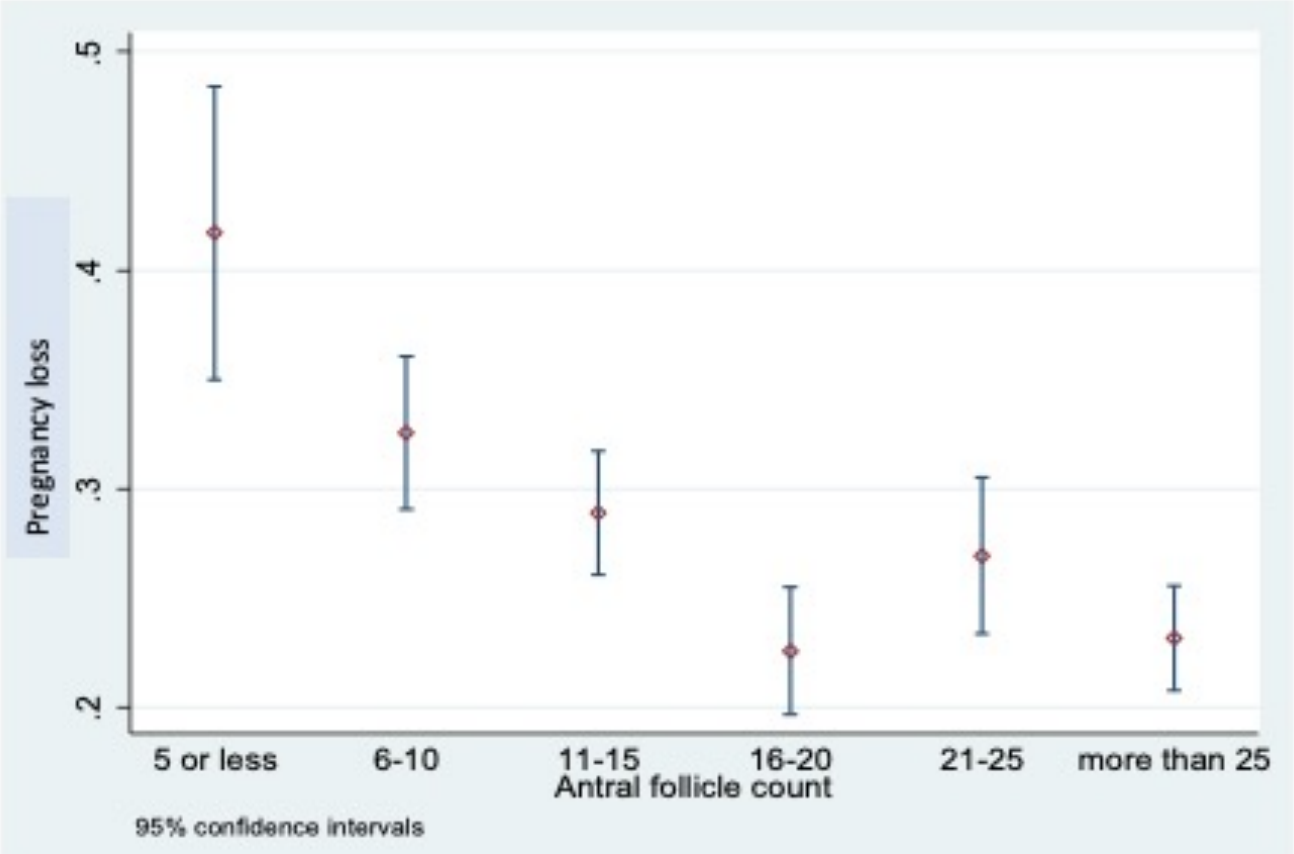


Figure 12: Pregnancy loss per AFC category (X axis shows AFC categories in increments of five follicles and Y axis shows the probability of pregnancy loss)

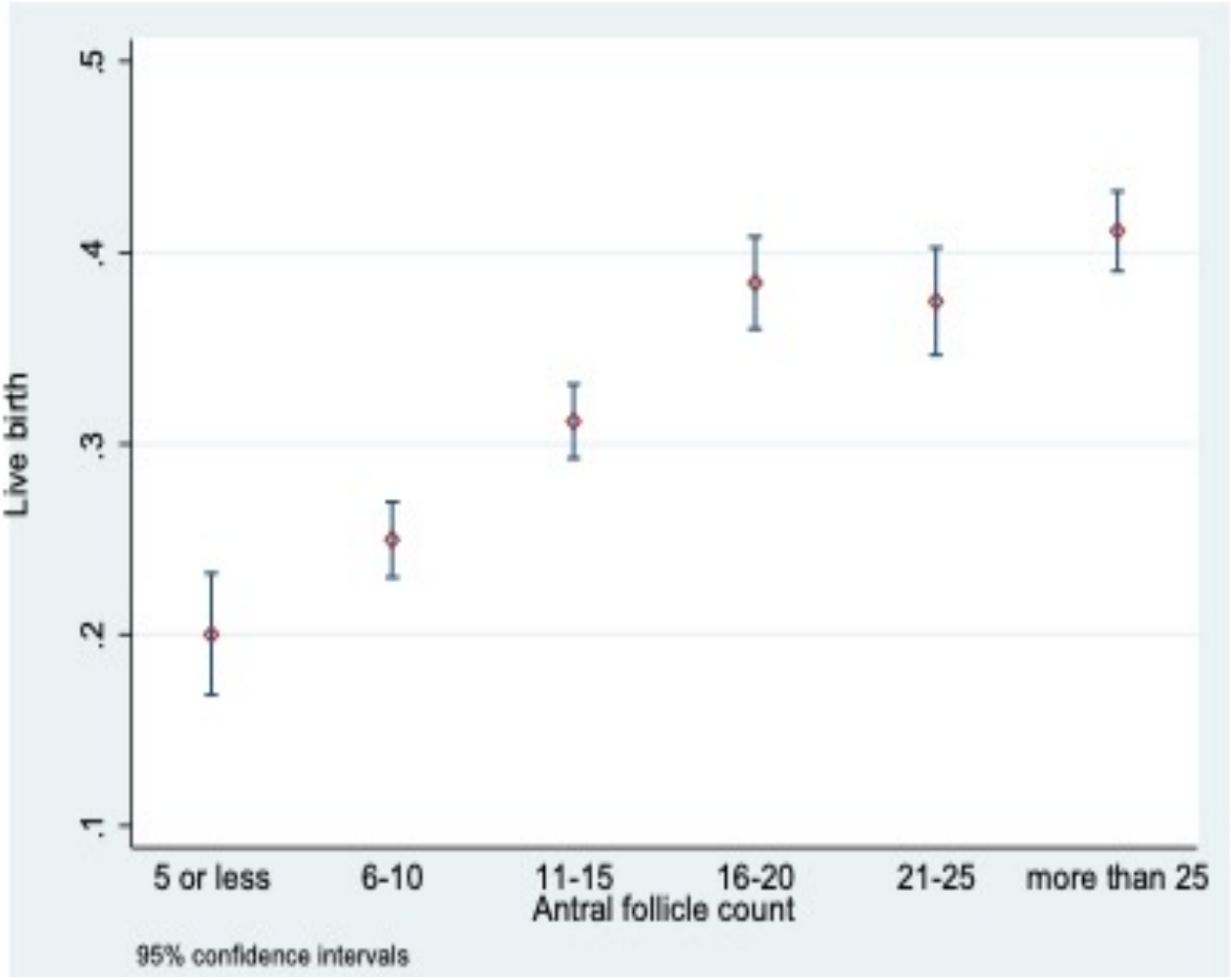


Figure 13: Live birth per AFC category (X axis shows AFC categories in increments of five follicles and Y axis shows the probability of live birth)

This association was independent of age, body mass index and ethnicity for both live births (crude RR 0.83; 95% CI 0.77—0.89 and adjusted RR 0.863; 95% CI 0.85—0.92) and pregnancy losses (crude RR 0.83; 95% CI 0.77—0.89 and adjusted RR 0.86; 95% CI 0.85—0.92).

When split into clinical and biochemical pregnancy loss, AFC categories were shown to have a stronger association with biochemical pregnancy loss in comparison to clinical pregnancy loss (Figure 14).

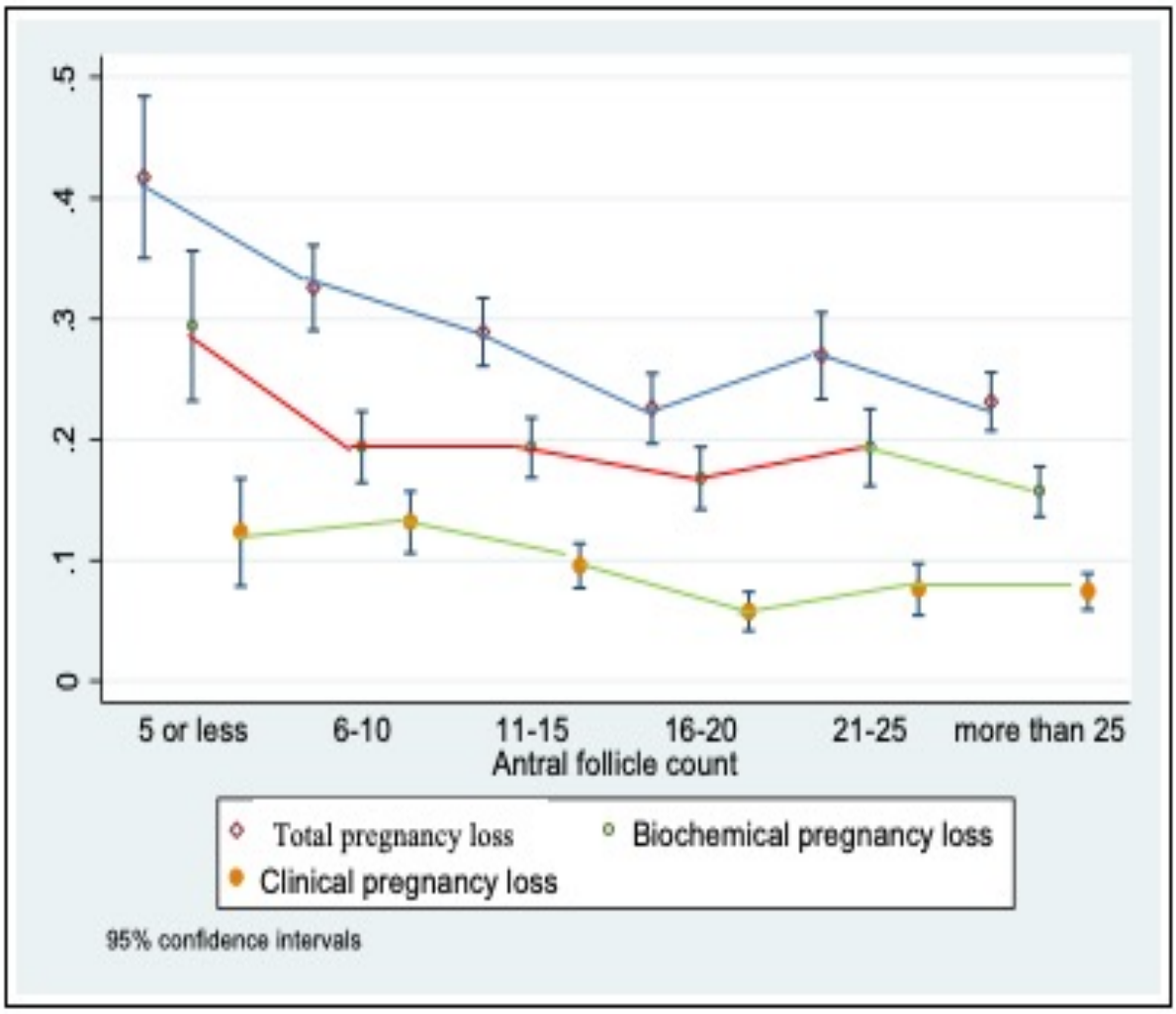


Figure 14: Biochemical, clinical and total pregnancy loss by AFC categories. (X axis shows AFC categories in increments of five follicles and Y axis shows the probability of pregnancy loss)

When we examined the impact of a low AFC (0-5 follicles) for young women of 20-30 years, where the pregnancy loss rate is expected to be low, it was found that the risk of pregnancy loss was almost double compared to women with a normal AFC (16-20) (RR 1.99; 95% CI 1.19-3.35; p=0.009).

Discussion

This study establishes an association between AFC and live birth rate and pregnancy loss, independent of age, ethnicity and BMI in patients having IVF/ICSI treatment. This is one of the largest studies carried out evaluating the association between AFC and pregnancy loss to our knowledge. As all IVF cycles that

were included within this study were of patients receiving treatment from the same provider, the extent of inter provider treatment variation is minimised.

However, this study does have a number of limitations. All women included in this study self funded their treatment. In UK, there are strict eligibility rules on access to state funded treatment, with older, multiparous and women with higher BMI being excluded. Therefore it is conceivable that there could be differences in the baseline characteristics and socio-economic status between women who were treated at CARE and women who receive state funded IVF, which could limit the generalisability of our study. As all data is from IVF cycles, it is difficult to draw conclusions about the relationship between AFC and pregnancy loss in women who conceive naturally. As frozen cycles and intra uterine inseminations were excluded from the analysis, these results are not applicable to all women having assisted conception. At CARE, AFC was not always measured in every patient, and is carried out at the clinician's discretion. Furthermore, there were no guidelines on how clinicians use AFC in determining the most appropriate treatment option. It is likely that AFC would have been taken into consideration by the clinicians when determining the IVF protocol and dose of stimulation medication, therefore the impact of this on the outcomes cannot be excluded. A future survey of clinicians based at CARE, establishing how AFC results affects their decision of which treatment protocol and medication dosage to use could enable us to make some assumptions in this area.

Inherent errors associated with the measurement of AFC, including inter observer variability, are also not accounted for. It is also plausible that some factors may have an impact on both the ovarian reserve (170) and the endometrium (171) independently. For example, smoking is shown to impact on the ovarian reserve as well as the endometrium. Therefore, the association between AFC and pregnancy loss may be accounted for in other ways.

We used Poisson regression to estimate the RR. This model enables one to calculate adjusted RR, which is deemed intuitive and preferred by many clinicians(168)(172). However for count data, such as ours,

poisson may not always be the most appropriate model as it is based on the assumption that the mean is identical to the variance(173). However the impact of any such errors on large datasets is limited. We were able to demonstrate the reproducibility of our results by carrying out binomial regression using generalised linear model to calculate ORs and RR for unadjusted live birth and early and late pregnancy loss. This has been listed in appendix 5. The similarity of results produced by both models gives confidence of the existence of the relationship between AFC and live birth and pregnancy loss, independent of the regression model used.

This study's finding showing association between AFC and live birth rate supports the findings of other studies, including Holte et al., and Fisch et al (78,174). Holte et al., is a large prospective cohort study, where IVF outcomes were adjusted to both the woman's age and the number of previous treatment. Holte et al., demonstrated a linear relationship between AFC and live birth rate up to an AFC of 30. Fisch et al., showed that those with an AFC >20 had a statistically significant higher live birth rate in comparison to those with an AFC<10. Interestingly, there was little difference in the live birth rate between the AFC categories of <10 and 10-19 follicles.

However our study's finding showing an association between AFC and live birth rate is in contrast to the findings of Hsu et al., and Li et al (157,162). Hsu et al., found that whilst AFC was predictive of ovarian response, it did not correlate with pregnancy rates, implantation rates or live birth rates. Whilst the investigators of this study had adjusted their results for age, they had not adjusted for ethnicity and BMI, despite the presence of statistically significant uneven distribution of these characteristics between AFC categories. It is difficult to ascertain if the differences between our study and this study could be accounted by this. Li et al., demonstrated that there was a trend indicating women who achieved live birth had higher AFC. However Li et al., did not find a statistically significant association between AFC and live birth rate. Li et al., did not adjust the analysis of AFC and live birth rate to possible confounding factors such as age, ethnicity and BMI. Further comment on whether the differences between Li et al., and our study could be accounted by this lack of factoring confounding variables cannot be made.

This study's finding of an association between AFC and pregnancy loss supports the findings of Keane et a (139)l, Mustafa et al (140)and Sahu et al (175). However this does contradict the findings of Bishop et al (131) and Haadsma et al (163). Bishop et al (131), studied the association between AFC and pregnancy loss by allocating patients into categories of low, normal and high AFC. Low was determined as an AFC count of 10 or below. However many clinical guidelines, including the NICE guidelines (8) set the threshold for a patient to be deemed to be having a low ovarian reserve to be at a value of four or below. Therefore by setting the bar of AFC to be at 10, Bishop et al., might have diluted the reporting of any relationship between AFC and miscarriage. Haadsma et al., was a prospective cohort study, which looked at both spontaneous conception as well as assisted reproduction. Therefore the study population was not comparable to our study.

There are a number of possible explanations for the relationship between AFC and live birth and pregnancy loss. The limited pool hypothesis explains that as those with low AFC are likely to have a lower number of follicles and yield less oocytes, the number of embryos available to pick for implantation is limited. Therefore it is possible that clinicians are forced to use poorer quality embryos in those with low AFC.

Work by Grande et al (176), has shown strong association between low AFC and rates of aneuploidy. It is possible AFC is a good surrogate marker of DNA damage of the oocytes. As vast majority of early pregnancy losses are due to chromosomal errors (Jacobs PA, 1987), it is plausible that the higher incidence of pregnancy loss in those with low AFC could be attributed to the higher incidence of chromosomal errors.

The implications of this study are numerous. Firstly, in an IVF setting, when encountering patients with a low AFC, clinicians should not restrict their discussion to the probability of IVF success, but also counsel them about the higher pregnancy loss rate. This is particularly important in younger women, who may not be expecting a high chance of having a pregnancy loss after implantation. Most literature currently focuses on age as the key predictor of the chances of IVF success and pregnancy loss. AFC is

an independent predictor of live birth and pregnancy loss, and should be used in addition to age to accurately counsel women on their individualised risk.

Further research is needed to determine whether a relationship between AFC and pregnancy loss following natural conception exists.

Conclusion

A relationship between low AFC and live birth and pregnancy loss, independent of age, ethnicity and BMI, in patients having IVF treatment exists. Further research is needed to explore if such relationship exists for women who have spontaneous conception.

Chapter five: inter-cycle variation in antral follicle count, anti-mullerian hormone and follicle stimulating hormone in women with risk factors for poor ovarian response- a prospective cohort study

Preamble to chapter five

Thus far in the thesis, I have established:

1. There is some month-to-month variability in ovarian reserve tests as measured by FSH, AMH and AFC. The variation appeared to be greater for FSH and AFC when compared with AMH, but the available data were limited and heterogenous in nature, preventing a meta-analysis.

2. There is an association between AFC and egg count. There is published evidence that egg count closely relates to live birth outcome.

3. There is some association between AFC, AMH, FSH and pregnancy loss.

However, due to the heterogenous nature of the existing studies, I could not draw firm conclusions about the degree of menstrual cycle to menstrual cycle variability, necessitating a primary study. Furthermore, existing studies did not specifically examine women with low ovarian reserve. As explained in chapter one, this is the group for which a small increase or decrease in the number of eggs collected is likely to have the greatest clinical impact.

To establish whether a menstrual cycle to menstrual cycle variation exists in ovarian reserve tests in women with low ovarian reserve, I carried out a primary study. This study and its results are described in this chapter.

Contributions

Dr Bala Karunakaran- Conceived the idea, carried out AFC and blood tests, data collection, data analysis and wrote this manuscript.

Dr Abey Eapen carried out some of the AFC measurements and blood tests and helped with quality assurance.

Mr Aurelio Tobias and Dr Adam Deval helped analysis using STATA and PRISM packages respectively.

Prof Coomarasamy proof-read the manuscript and provided substantial edits.

Abstract

STUDY QUESTION: Is there variation in the ovarian reserve markers of AFC, AMH and FSH from one menstrual cycle to another, often referred to as inter-cycle variation, in women with low ovarian reserve?

SUMMARY ANSWER: This study proves a menstrual cycle to menstrual cycle variation exists for AFC, AMH and FSH in women with low ovarian reserve.

WHAT IS ALREADY KNOWN: Multiple studies had commented on the inter-cycle variation of ovarian reserve markers, though very few studies had made direct comparison with each other. There are no studies that have studied this variation in women with low ovarian reserve.

STUDY DESIGN, SIZE, AND DURATION: A prospective cohort study of 47 women over three consecutive menstrual cycles

PARTICIPANTS/MATERIALS, SETTING, METHOD: 47 women with an identified risk factor for low ovarian reserve, recruited from a large tertiary fertility clinic (Birmingham women and children's NHS Trust) in United Kingdom.

MAIN RESULTS AND THE ROLE OF CHANCE: The mean of AFC Max variation, AMH Max variation and FSH Max variation were 3.87, 9.61pmol/L and 6.48 respectively. When analysis was restricted to women who we deemed to have low ovarian reserve(AFC of 10 or less, AMH of 5.5pmol/L or less or FSH of 8.9IU/ml or more for one or more measurements), the mean of AFC Max variation, AMH Max variation and FSH Max variation were 2.4, 1.2pmol/L and 3.8IU respectively.

LIMITATIONS AND REASONS FOR CAUTION: Whilst every effort was made to reduce measurement errors, inter and intra operator variability and its potential impact on the results cannot be entirely eliminated. The clinical impact of this observed inter-cycle variability is also unknown.

WIDER IMPLICATIONS OF THE FINDINGS: The existence of inter-cycle variation means clinicians should be cautious in forming judgements about a patient's ovarian reserve or whether a patient

should be treated or not based on a single measurement Further research is needed into the clinical implications of this inter-cycle variation, including whether starting IVF treatment on a month with a test showing high ovarian reserve would result in better reproductive outcomes.

TRIAL REGISTRATION: researchregistry3201

Introduction

Ovarian reserve tests are routinely carried out by clinicians prior to starting assisted conception treatments (100). Clinicians use these tests as aids to determine the suitability for IVF or ICSI, the best treatment protocol, the dose of medication used for ovarian stimulation and in counselling the couple about IVF success (177).

There are a plethora of ovarian reserve tests available. NICE guidelines recommend that clinicians restrict their investigations to testing for antral follicle count (AFC), anti-mullerian hormone (AMH) levels and follicle stimulating hormone (FSH) levels, as these tests have the strongest evidence supporting them (8).

The relationship between ovarian reserve tests and IVF outcomes have been demonstrated in multiple studies (44). The analysis carried out in chapter three showed a positive relationship between AFC and live birth rate (LBR). The analysis in chapter three also demonstrated a decrease in pregnancy loss rates as the AFC increased. Therefore it can be expected that higher the ovarian reserve, the more likely a woman is to have a successful IVF outcome. There are multiple reasons why low ovarian reserve may have such impact. Firstly, those with poor ovarian reserve are more likely to have a poor response to controlled ovarian hyperstimulation (178). The consequence of this could either be cancelling the IVF cycle or having a lower number of oocytes extracted, with less chance of fertilisation or development of a good quality embryo (179). The increased pregnancy loss associated with low ovarian reserve has two potential explanations. The limited pool hypothesis, as outlined in chapter three, suggests that as the number of oocytes is limited, the number of embryos that are available to choose from also becomes limited, resulting in poorer quality embryos being used and therefore higher rates of pregnancy loss. Some academics have suggested that ovarian reserve test values may reflect the ageing process of the ovary and that low values may indicate ovaries which have undergone accelerated ageing, non-concordant with the actual age of the woman (180). As older women have higher rates of miscarriage and embryos with chromosomal abnormalities, women with abnormally aged ovaries could also have higher rates of pregnancy loss (181)(172).

The systematic review in chapter two suggests that there is month to month variation in ovarian reserve tests within the same individual. However as also described in chapter two, included studies looked at either healthy volunteers or women with subfertility, and did not specifically focus on women with low ovarian reserve. Cycle to cycle variation in a woman with normal or high ovarian reserve is unlikely to be clinically important. The work by Sunkara et al., showed a positive correlation between the number of eggs retrieved, from one to 15 eggs and live birth rate (23). Beyond 15 eggs, no correlation was found.

The work by Hsu et al., has shown strong correlation between AFC and the number of oocytes retrieved (162). For women with a high AFC, month to month variation is unlikely to have an impact as they would have a high yield of oocytes resulting in a high quality of embryos to choose irrespective of the month. However, for a woman with an AFC of two in one month and six in another month, this difference could be clinically important.

There is a need to investigate whether menstrual cycle to menstrual cycle variation exists within individuals who have a low ovarian reserve and thus likely to have a poor ovarian response. The aim of this prospective cohort study was to establish the inter-cycle variation in AFC, AMH and FSH in women who were at risk of having a low ovarian reserve.

Measuring inter-cycle variation

There is little consensus on how to measure variation of a test result within the same individual. Studies looking at inter-cycle variation have deployed various statistical tools to study this. Some authors have calculated inter-cycle variation as mean difference (104,105,107,108). Mean difference is derived by calculating the difference between two observed values and subsequently calculating the average of all the differences (182). This approach has considerable limitations. As the difference between the first and second observation could be either in positive or negative magnitude, when the differences are added together they are likely to cancel each other. Therefore mean difference which is an average of the values will be small and underestimate the true variation.

Fanchin et al., calculated the inter-cycle variation as intra-class coefficient (ICC) (51). Intra-class coefficient was first described by Shrout and Fleiss, and is now a widely used tool to assess the reproducibility of an investigation (183,184). McGraw defined intra-class coefficient as the proportion of variance as attributable to the measurement (185). However intra-class coefficient is limited by studying the overall agreement between the two cohort measurements; therefore it is not best placed to study the agreement between multiple measures from the same individual.

Some studies, including Brown et al., and Rastamov et al., used coefficient of variation to measure inter-cycle variation (103,104). Coefficient of variation is the difference between the values in relation to the mean of the values (109). As coefficient of variation takes into account the mean, the resultant value is skewed by the value of the mean and thus limited in its generalisability.

Due to these limitations, many authors, including Jayaprakasan et al., have used Bland Altman plots and limits of agreement to express variation (62). A Bland Altman plot graphically displays the difference between two observations plotted against the average of the two measurements (186). Limits of agreement are calculated as the mean plus and minus 1.96 times the standard deviation (92). Limits of agreement are drawn as a line on the graph. This method has the advantage of studying the difference between the values specifically. It also allows values of positive and negative magnitude to be plotted on the same graph. As the mean and standard deviation is calculated based on every value that has been entered, the limits of agreement represent the entire dataset and have limited value in assessing an individuals' inter-cycle variation. Despite this, a graphical representation is useful for identifying outliers with the greatest inter-cycle variation.

Methods

Study Design

Before beginning the cohort study, the aim was to recruit 47 women who were being investigated for subfertility. The assumption of 10% drop out after commencing investigation was anticipated. The sample size was determined as previous studies had been able to establish inter-cycle variation for AFC,

AMH and FSH with between 40-50 paired measurements (a summary of previous studies can be found in Chapter two of this thesis (**Table 3**). The sample size was also limited for pragmatic reasons based on the funding available to carry out AMH and FSH tests and the number of scan slots available to perform AFC.

The study was approved by the North of Scotland Research Ethics Service (16/NS/0104 IRAS 204528). Women were recruited from secondary care fertility outpatient clinics at the Birmingham Women's and Children's Hospital (BWCH).

The inclusion criteria were designed to identify women at risk of having low ovarian reserve. According to the Human Fertilisation and Embryology Authority (HFEA) October 2007 data, which were also featured in the 2013 NICE Guidelines (shown in appendix), live birth rate per embryo transfer starts to decline steeply after the age of 35. Therefore women aged over the age of 35 were included in the study. The NICE guideline designates an FSH level of greater than 8.9 IU/ml or an AMH of less than 5.4 pmol/L as abnormal (8). Hence, these values were adopted for AMH and FSH. As AFC was not measured at the outpatient clinic, no inclusion criteria based on AFC was adopted. Any woman with a previous history of poor response to controlled ovarian hyperstimulation was also included in this study. We included women with one of the risk factors for low ovarian reserve in this study.

Prospective participants were identified by the secondary care clinicians and introduced to the researcher. A participant information sheet was provided(appendix 2), and prospective participants were invited to discuss this further with the researcher. In instances where the prospective participant was not fluent in English, an interpreter was provided. At the end of the discussion, if the prospective participant was willing to proceed, written consent was obtained(appendix 3). Participants were given the contact details for a dedicated telephone number for this study, and instructed to make contact as soon as their next menstrual cycle commenced.

Participants were invited to attend during their early follicular phase, (defined for the purpose of this study as days two to six of their menstrual cycle) for three consecutive menstrual cycles. The decision

for three consecutive menstrual cycles was taken as this time period fell within the average time women waited between their initial and follow up visit at this hospital. Therefore our research was designed not to delay the commence meant of IVF treatment. The participants' understanding of the study and their consent to proceed was then reconfirmed by the researcher before taking the first measurements. A venepuncture was performed and blood was collected using an Eclipse® needle and vacutainer system with Serum Selective II Tubes (SSTII). SST II tubes have been shown to be superior to the alternative Serum Selective Plus (SST) tubes in studies, with less volume dependent variation in results (187). Each of the participants was tested for AMH and FSH levels. The analysis of serum was carried out using electrochemicaluminescence on a Roche Cobas e801™ platform (188). All serum samples were carried out by an accredited laboratory (TDL laboratories London).

All participants had a transvaginal ultrasound scan performed with consent and a chaperone. All scans were performed by one of two investigators to keep inter-observer variability to a minimum. 20 scans had both investigators scanning each participant together to ensure agreement and quality assurance. Participants were asked to empty their bladders and then subsequently placed in a modified Lloyd Davies position. All scans were performed using a Voluson S8, GE Healthcare ™ Scanner with a 3.8-9.3MHz endocavity volume probe. A baseline pelvic anatomy scan was carried out to assess for any pelvic pathology prior to performing an antral follicle count.

For this study, follicles measuring 2-10mm were recorded as antral follicles (153). Real-time two dimensional ultrasound scan (2D US) was used. Participants were scanned in two planes (longitudinal and transverse) to identify the optimal plane for identifying and counting the antral follicles. Each ovary was scanned from one ovarian margin to another, using the technique described in the consensus opinion published by Coelto Neto et al. (189). Ultrasound scan was performed on each ovary and antral follicles measured and recorded separately.

Statistical analysis

To establish the greatest variation within one individual, for each ovarian reserve test, the lowest obtained value (trough result) of the three visits was subtracted from the highest value (peak result) of the three

visits, to produce the maximum variation. The mean of the maximum variation and its standard deviation was calculated. Further descriptive statistical analysis, such as bland-altman plots, limits of agreements, medians, inter-quartile ranges were calculated.

Most statistical calculations were done using STATA version 14 (Stata corp, College station, TX, USA). Bland-altman plots were produced using PRISM version 7 (Graphpad, San Diego, USA).

Results

Baseline characteristics

The baseline characteristics of participants are displayed in Table 13. The majority of the women were aged over 35 (88.9%), non-smokers (93.2%) with high FSH levels (55.6%), primary infertility (75.5%) and no previous fertility treatment history (93.3%).

Table 13 Baseline characteristics

Average age in years(SD)	39.2(2.9)
Ethnicity	
White(%)	18(40.0)
Black(%)	2(4.4)
Chinese(%)	3(7.0)
South Asian(%)	15(33.3)
Other(%)	7(15.5)
Smoking status	
Yes(%)	2(4.4)
No(%)	30(66.6)
Previous smoker(%)	12(26.6)
Fertility status	
Primary infertility(%)	34(75.5)
Secondary infertility(%)	11(24.4)
Previous IVF(%)	3(6.7)
No Previous IVF(%)	42(93.3)
Ovarian reserve status	
Age >35(%)	40(88.9)
Low AFC(%)	8(17.8)
Average of AFC (SD)	8.3(5.7)
Low AMH(%)	18(4.0)
Average of AMH	9.2(10.0)
High FSH(%)	25(55.6)
Average of FSH(SD)	10.9(7.2)
Previous poor response (%)	1(2.2)

Pattern observed over three visits

Graphical displays of the results of all visits by participants, and results for those with low AFC, low AMH and high FSH are shown in the figures below.

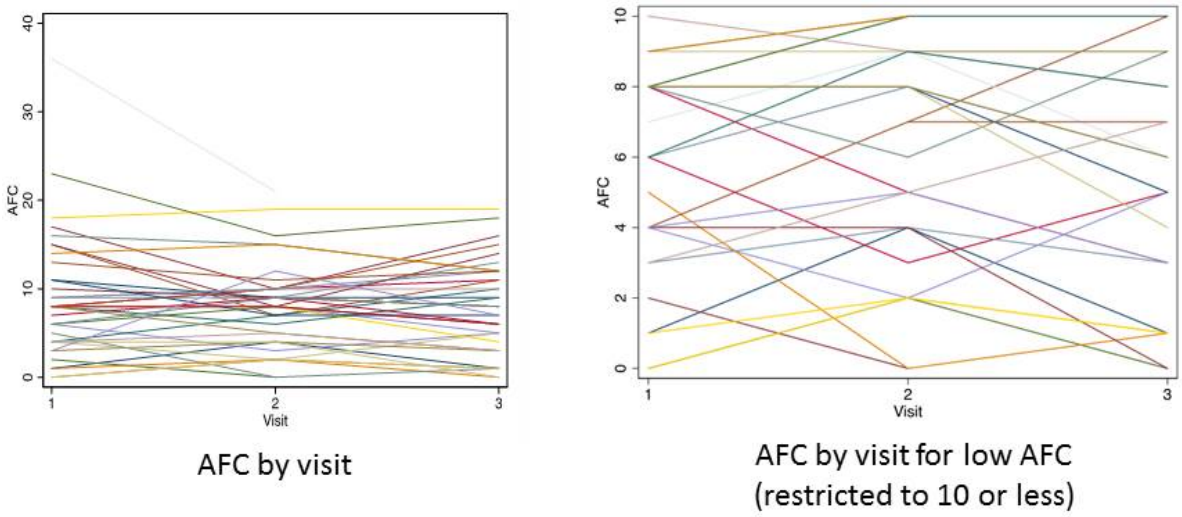


Figure 15: AFC value obtained plotted against each visit

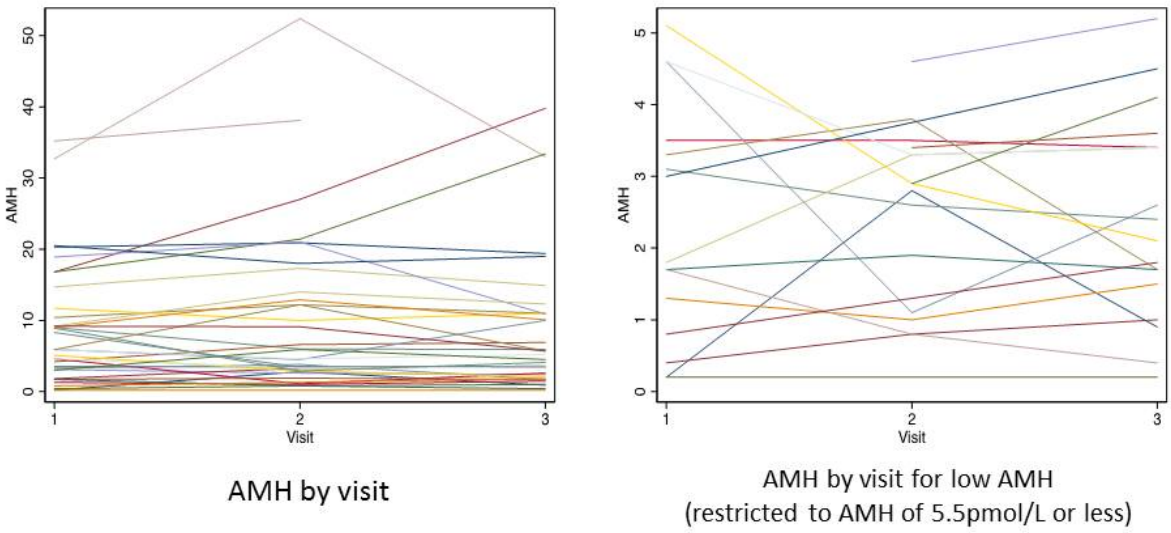
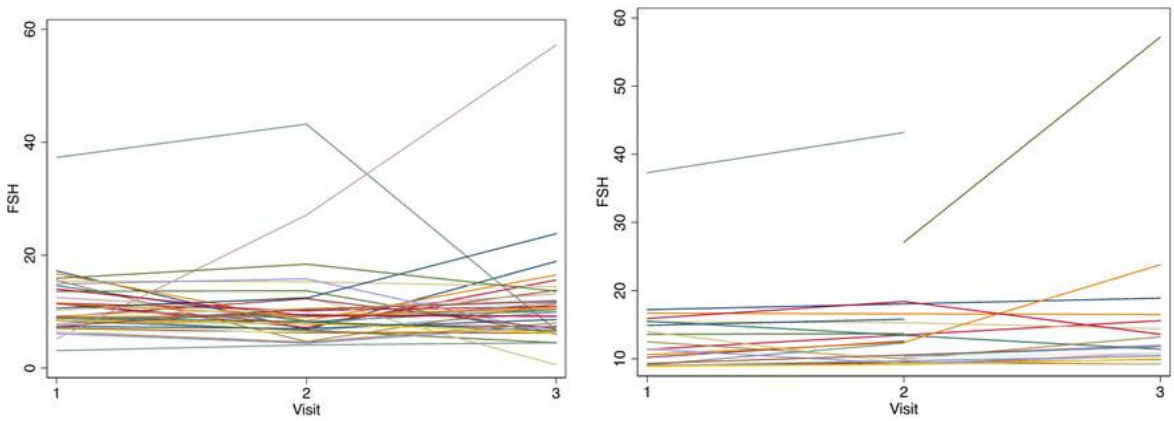


Figure 16: AMH value obtained plotted against each visit



FSH by visit

FSH by visit for high FSH
(FSH 8.9IU or higher)

Figure 17: FSH value obtained plotted against each visit

Table 14 shows the summary characteristics of the observations from the three ovarian reserve tests amongst the women studied. For each study the mean difference between month one and month two and month two and month three were calculated. The largest magnitude of observed difference within each participant was labelled AFC max variation, AMH max variation and FSH max variation respectively. This was calculated by subtracting the highest value obtained in all measurements from the lowest value obtained from all measurements. Means, standard deviation and interquartile range of the maximum variation values are displayed in Table 15.

Maximum variation in women with lower ovarian reserve, defined as an AFC of 10 or less, an AMH of 5.5pmol/L or less and an FSH of 8.9IU or higher, in one or more measurements, were analysed separately. Means, standard deviation and interquartile range of the maximum variation values are displayed in Table 15.

Table 14 Mean of differences, standard deviation, upper limits of agreement and lower limits of agreement

ORT	Mean of differences	Standard deviation (SD)	Upper limits of agreement (ULA)	Lower limits of agreement (LLA)
AFC				
Month1-Month2	0.95	4.07	8.93	-7.03
Month2-Month3	0	2.92	5.72	-5.72
Difference between highest AFC measurement obtained and the lowest AFC measurement obtained (AFC Max Variation)	3.87	2.75	1.51-	9.25
AMH				
Month1-Month2	-1.89	4.58	7.79	-10.16
Month2-Month3	0.58	5.29	10.9	9.79
Difference between highest AMH measurement obtained and the lowest AMH measurement (AMH Max variation)	9.61	8.06	6.18	-25.40
FSH				
Month1-Month2	-9.61	8.06	6.18	-25.40
Month2-Month3	-0.79	9.47	17.78	-19.36
Difference between highest FSH measurement obtained and the lowest AFC measurement obtained (FSH Max variation)	6.48	9.74	12.61	-25.57

Table 15 Maximum variation of AFC, AMH and FSH

	Median	Quartile one	Quartile three	Interquartile Range
AFC Max	3.0	2.0	5.0	3.0
AMH Max	7.8	2.2	15.0	12.8
FSH Max	3.8	1.6	8.0	6.4

Table 16 Maximum variation in women with low AFC, low AMH and low FSH

	Mean	Standard deviation	Median	Quartile onw	Quartile three	Interquartile Range
AFC Max in low AFC	2.4	1.5	2.0	2.0	3.0	1.0
AMH Max in low AMH	1.2	1.0	1.1	0.5	1.6	1.1
FSH Max in low FSH	3.8	6.1	1.8	1.1	4.1	3.0

Bland-Altman plots were created for the difference between visit one and visit two, visit two and visit three, visit three and visit one, as well as for the maximum difference (the difference between the highest and lowest measurement).

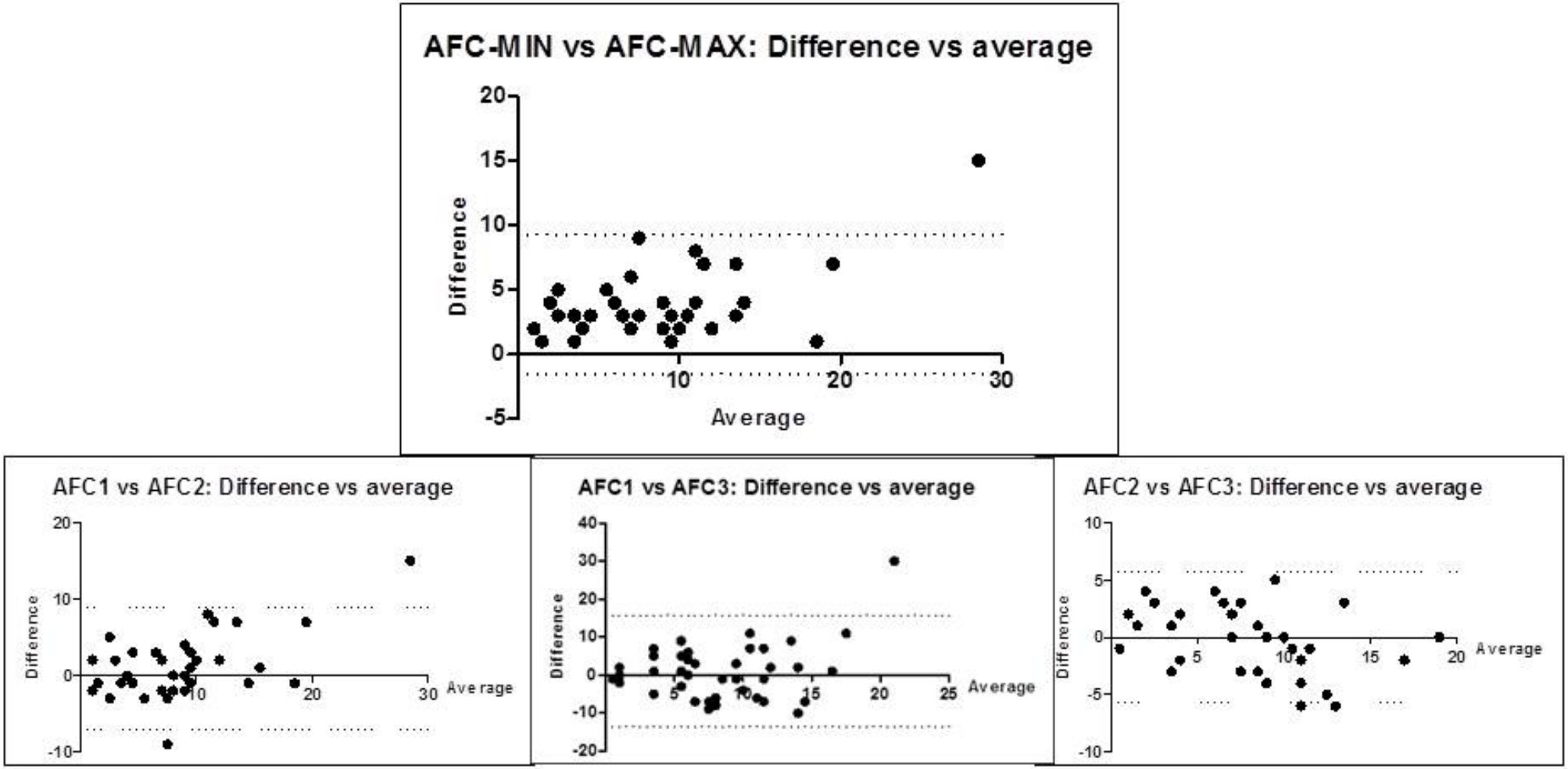


Figure 18: AFC bland-altman plots- difference between the values (y axis) plotted against the average of the two values (x axis).

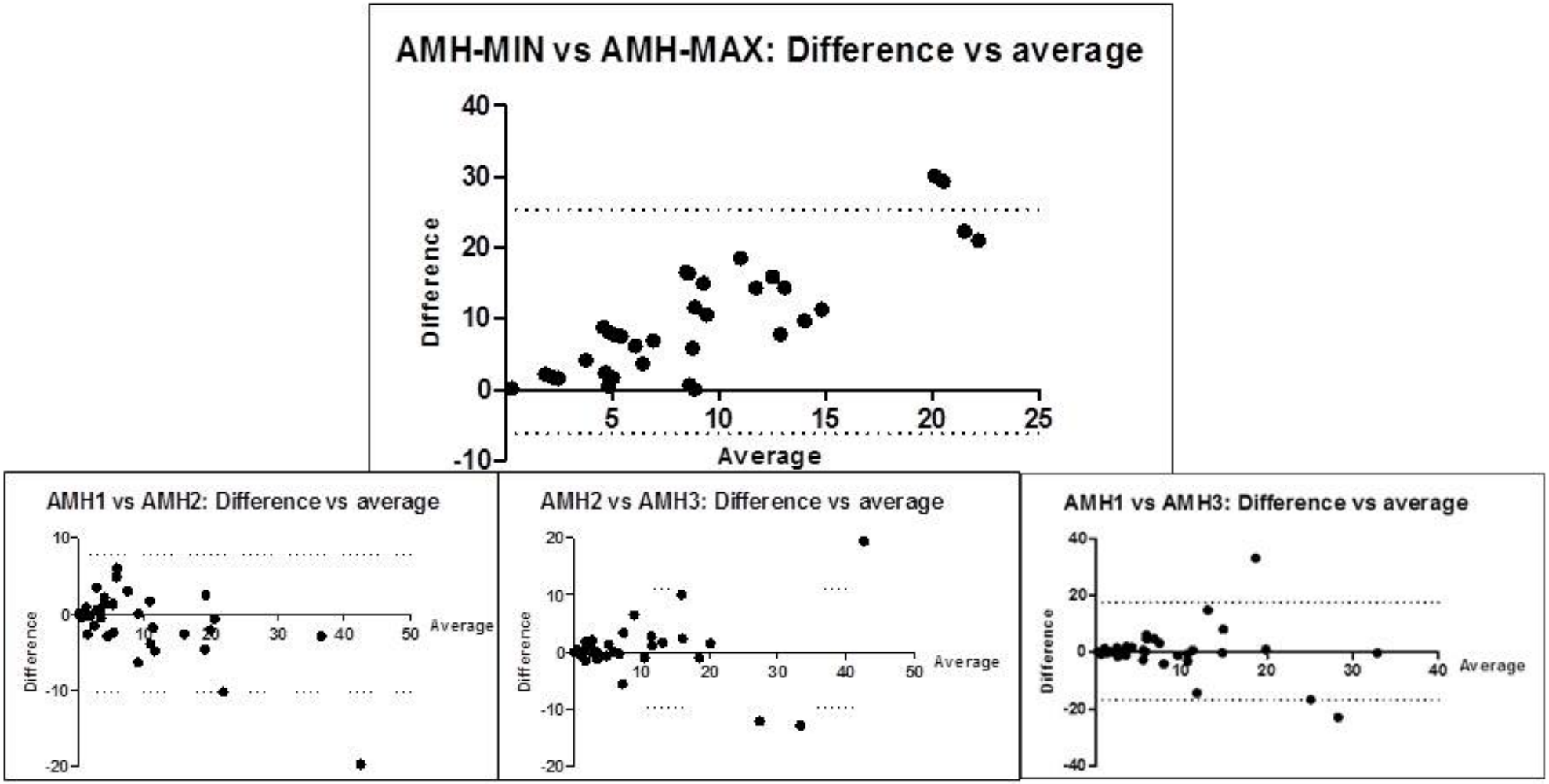


Figure 19: AMH bland-altman plots- difference between the values (y axis) plotted against the average of the two values (x axis).

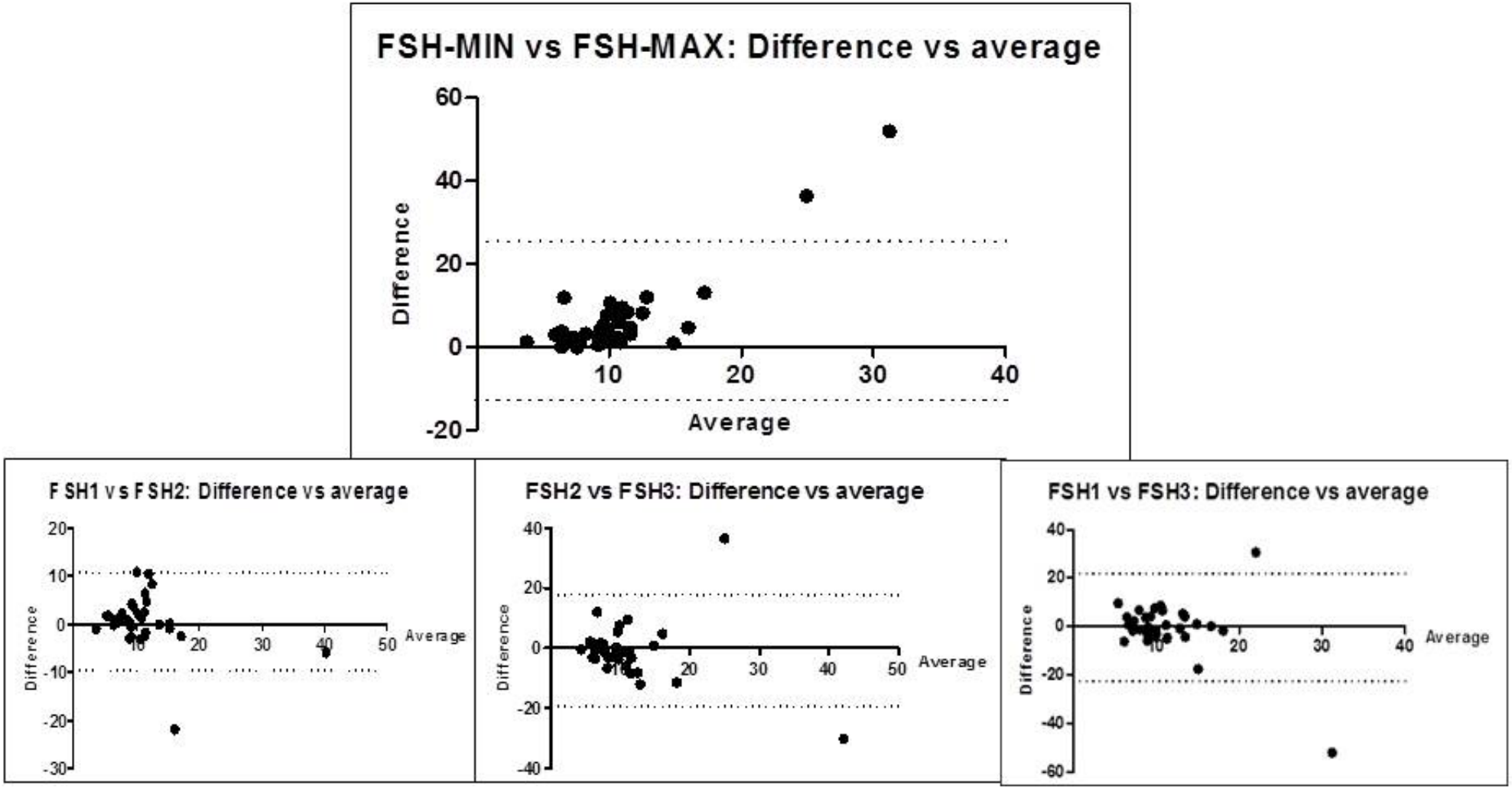


Figure 20: FSH bland-altman plot- difference between the values (y axis) plotted against the average of the two values (x axis).

Discussion

This study proves a menstrual cycle to menstrual cycle variation exists for AFC, AMH and FSH in women with low ovarian reserve. AFC Max variation, AMH Max variation and FSH Max variation, which were calculated by subtracting the lowest obtained value from the highest obtained value within each individual, is the measurement of most importance. This is because as this value reflects the highest variation between the three visits, it describes the biggest potential menstrual cycle to menstrual cycle variation within any individual. The mean of AFC Max variation, AMH Max variation and FSH Max variation was 3.87, 9.61pmol/L and 6.48IU/ml respectively. When analysis was restricted to women who we deemed to have low ovarian reserve (AFC of 10 or less, AMH of 5.5pmol/L or less or FSH of 8.9IU/ml or more for one or more measurements), the mean of AFC Max variation, AMH Max variation and FSH Max variation was 2.4, 1.2pmol/L and 3.8IU/ml respectively.

Study strengths

There are several strengths to this study. This is one of the few studies which looked at inter-cycle variation of ovarian reserve tests as the primary outcome. This is the only study to our knowledge which has specifically focused on women with an identified risk factor for poor ovarian response to controlled ovarian hyperstimulation. Therefore the findings of this study are generalisable to the population at risk of poor ovarian response to controlled ovarian hyperstimulation. Birmingham is an ethnically diverse, multicultural city and its residents comprise of a wide range of socio-economic backgrounds. Consequently, this study is likely to capture the diversity of the UK population. Unlike some of the other studies which have analysed data retrospectively or used data collected for routine clinical use, this study collected data prospectively. The ultrasound machine and settings were standardised and the same laboratory assays were used for all participants, reducing any systematic errors in measurements. All scans were performed by one of two investigators, keeping variation between operators to a minimum and reducing inter observer variation. As the three scans and blood tests were carried out in three consecutive menstrual cycles, the resultant variation due to age related decline in ovarian reserve is minimised.

Study limitations

There is some debate as to whether other techniques, such as cine-loop ultrasound scan, three dimensional ultrasound scan or sono AVC may be more accurate in measuring AFC (189). Sono AVC uses computer automation to measure and count the follicles and can reduce the uncertainty posed by human judgement (98). However it has been shown that automatic evaluation of follicles can often count para-ovarian structures as follicles (189). Cine-loop ultrasound scan has the benefit of allowing the scan to be recorded and played back and therefore to have an independent verification by a second person (190). Three dimensional mode allows the investigator to look at the ovaries in multiple planes. Advanced functions such as rendering mode which enhances the contrast, and inversion mode which allows complete visualisation of all follicles within the required volume are also available in three dimensional mode—these can make the AFC more accurate (191). However the consensus opinion by the International Society of Ultrasound in Obstetrics & Gynecology (ISUOG), stated that there is no clear evidence that one technique was superior to another (189). Furthermore, these techniques require advanced training, expensive equipment and software, and are not routinely performed by care providers. As one of the aims of this study is to lead to results that could be clinically relevant and reproducible, a decision was made to use conventional two dimensional real-time scanning.

There were some limitations due to study design. A similar study with a bigger sample size and one that was multi-centred could provide greater confidence in generalisability. For pragmatic reasons the number of cycles were limited to three in this study. Continuing to observe women over a greater number of consecutive cycles could be informative.

Investigators were not blinded to results of previous month's scans. It was decided that due to the small number of participants in the study at any given timeframe, an investigator recollecting the scan performed the previous month could not be avoided. Therefore the bias introduced by the knowledge of the previous results cannot be eliminated. As all ultrasound scans had to be performed during the early follicular phase of the menstrual cycle, there was a need for an investigator to be on standby throughout

the month. Therefore a pragmatic decision was taken to have all ultrasound scans performed by one of two investigators. Therefore the impact of inter-operator variability cannot be excluded.

Another weakness of this study was that whilst participants were chosen for having a risk factor present for reduced ovarian reserve, not all participants had truly reduced ovarian reserve. This has been partially accounted for in analysis, where calculations were restricted to those with proven reduced ovarian reserve tests.

Mechanisms of action

Chapter one of this thesis describes the hypothalomo-pituitary-ovarian axis and the interaction between AFC, AMH and FSH. FSH is produced by the pituitary gland. Studies have shown that environmental stress, nutritional state and body mass index have direct impact on the pituitary gland and subsequent FSH levels, therefore levels are transient (192,193). It has been reported that very small fluctuations weight can result in disruption to pituitary function (194). This could account for the inter-cycle variation observed in FSH in this study. FSH plays a crucial role in recruiting pre-antral follicles and in the process of pre-antral follicles developing into antral follicles (55). Hence AFC can reflect both the ovarian reserve of the woman as well as the levels of FSH in the preceding weeks. It could be argued that the same external factors which affect the FSH levels, could also affect the AFC. It could be suggested that external factors like prolonged stress or prolonged poor nutritional state may have a greater influence on AFC than transient changes. AMH is predominantly produced by preantral and small antral follicles, and not by large antral follicles (162). As the influence of external factors and FSH levels on pre-antral and small follicles is minimal, the influence of such factors on inter-cycle variation of AMH may be limited (195).

Comparison to existing studies

The results of this study are in agreement with many other studies that looked at inter-cycle variation, including Bancsi et al., Elter et al., and Hansen et al.,(107)(105)(196). It should be noted that Elter et al., also looked at all three ovarian reserve tests that were investigated in this study. None of the studies listed above investigated women particularly at risk of poor response to controlled ovarian hyperstimulation.

The results of this study are in contrast to the findings of Jayaprakasan et al., (62). Jayaprakasan et al., found that AFC had the least inter-cycle variation. Jayaprakasan et al., was a prospective cohort study which had the inter-cycle variation of ovarian reserve tests as the primary outcome. Jayaprakasan et al., also investigated AFC, AMH and FSH, however there were variations in the study design, which may account for some of the differences in findings. Jayaprakasan et al., investigated women with subfertility, but did not specifically analyse women with a risk factor for poor response to controlled ovarian hyperstimulation. Jayaprakasan et al., also took the measurements at consecutive IVF cycles, and not consecutive menstrual cycles. The study protocol specified that there had to be a minimum of two menstrual cycles between one IVF cycle and another, therefore there was a time delay between the two measurements. There are many studies which document the age related decline in ovarian reserve tests and the effect of time on the measurements cannot be excluded (63,114,197). Jayaprakasan et al. also restricted the investigation to two IVF cycles, unlike this study which studied three consecutive menstrual cycles. Therefore the likelihood of discovering inter-cycle variability is increased in this study.

Clinical implications

The establishment of inter-cycle variation in women with low ovarian reserve has clinical implications. The study by Moon et al., demonstrated that AFC positively correlated with the number of eggs collected, even in women with low AFC count (84). Sunkara et al., have demonstrated that eggs collected positively correlate with live birth rate, with a linear relationship as eggs collected increase from one to 15(23). Therefore it is likely that this menstrual cycle to menstrual cycle variation that we observed in women is of clinical importance.

Clinicians must also exercise caution in making clinical decisions or counselling women based on a single measurement of an ovarian reserve test. It is important that clinicians do not make important clinical decisions such as deciding not to proceed with IVF treatment based on a single result. In women with low ovarian reserve, AMH showed less Max variation in comparison to AFC and FSH. Therefore when interpreting results in women with low ovarian reserve, clinicians could interpret the AMH to be a

stable predictor of the woman's ovarian reserve, whilst using AFC and FSH as dynamic tests to assess the functional state of the ovary in that particular menstrual cycle.

Implications for future research

Women with poor ovarian reserve are known to have adverse IVF outcomes (44,177). As inter-cycle variation in ovarian reserve tests has been established in women with low ovarian reserve in this chapter, it poses the possibility of timing IVF treatment with optimal ovarian reserve test values. AFC is likely to be the most useful test, as this study has shown that it is likely to have clinically significant inter-cycle variation, even in women with low ovarian reserve. Further research, in the form of a clinical trial where IVF is started in the optimal cycle for the individual woman, may yield a new approach to IVF in women with low ovarian reserve.

We did not correlate the relationship between each ovarian reserve test value, taken at the same visit, within the same individual. Further research with a larger sample size, studying paired ovarian reserve tests and the relationship between them could help to identify the underlying biological mechanisms and better inform clinicians in interpreting them.

Chapter six: perception of diminished ovarian reserve: a qualitative study

Preamble to chapter six

So far in my thesis, I have shown that women with low ovarian reserve have poor reproductive outcomes. I have shown that ovarian reserve tests showed association with reproductive outcomes such as pregnancy loss and live birth rate. I have also demonstrated that AFC positively correlates with live birth and negatively correlates with pregnancy loss. A systematic review conducted as part of this thesis showed that there was menstrual cycle to menstrual cycle variation in ovarian reserve tests. The prospective cohort study I carried out showed that there is menstrual cycle to menstrual cycle variation in ovarian reserve tests in women with low ovarian reserve. There was a cycle to cycle maximum variation of 2.4 noted with AFC. The next step in this research journey would be to test the hypothesis that identifying a month with a high AFC in a particular woman, and starting treatment on that particular month, could lead to an improved reproductive outcome.

In this chapter, I describe the prospective qualitative study I carried out to understand the needs of women with low ovarian reserve and explore their ideas and expectations. I explore their attitude towards research and their level of acceptability in taking part in a research study as proposed above.

Contributions

Dr Bala Karunakaran- Conceived the idea, carried out interviews, transcribed, carried out analysis and wrote this manuscript.

Dr James Cheshire was the second coder and helped with analysis.

Prof Coomarasamy proof-read the manuscript and provided substantial edits.

Abstract

Study Question

What are the views of women with low ovarian reserve with regards to their fertility? How acceptable taking part in research is to women with low ovarian reserve?

Summary Answer

There is a lack of awareness of their ovarian reserve and potential for age related decline. This lack of awareness was more prevalent in certain cultural and social groupings. Women with low ovarian reserve wished that there had been more education on this topic at an earlier stage in their life, either at school or opportunistically by GPs. There is a strong desire to contribute to research amongst this population. Women were happy to accept a short delay in treatment if such delay could result in better IVF outcomes. Whilst the concept of a randomised controlled trial was not intuitively appealing, women were willing to participate in such trials if there was an overall compelling case made.

What is known already?

It is known that women are choosing to delay conception worldwide. There are published studies showing that the general population has poor awareness of the age related decline in fertility and overestimate the success of IVF treatments. There is little published about women with low ovarian reserve and their ideas and expectations, as well as their acceptability of clinical research.

Study size, design, and duration

This prospective qualitative study included 21 women over a six month period.

Participants/materials, setting, methods

This was a qualitative study which prospectively recruited women attending the fertility clinics of Birmingham women and children's hospital and had a risk factor for low ovarian reserve present. Purposive sampling with maximum variation was carried out to select participants. Semi-structured interviews were conducted and thematic analysis using framework method was carried out.

Main results and the role of chance

Results were presented in four areas; women's perception of ovarian reserve and their reasons for delaying starting family, need for further awareness of ovarian reserve, acceptability of an antral follicle count based IVF treatment protocol and women's attitude to research and acceptability of delaying treatment to identify the optimal cycle. Summary of the themes is presented above.

Limitations, reasons for caution

This study only recruited women from one fertility clinics. Therefore even with the use of maximum variation sampling, caution is needed about generalisability.

Wider implications of the findings

This research highlights the need for strategies to increase age related decline in ovarian reserve amongst women. Further research is needed to identify strategies to improve understanding in ethnic minority groups. When designing research for women with low ovarian reserve, considerations need to be made to ensure that any delay in starting treatment is kept to a minimum. If a randomised control trial design is used, investigators should be mindful of the reservation about this form of research amongst this population and should ensure that the rationale for such design is clearly communicated to potential participants.

Study funding/competing interest(s)

None

Introduction

Most women have an expectation that they will have children at some stage in their lives (198). It has been shown that women's perception of their chances of conception tend to be more optimistic than their actual chances of conception (199).The general understanding of fertility and factors affecting chances of fertility is poor, even amongst those who are seeking help to conceive (200).

Ovarian reserve is a term used to describe the number and quality of primordial follicles left in a woman's ovaries (201). Primordial follicles contain all the eggs that can be fertilised in a woman's lifespan (202). Ovarian reserve, low ovarian reserve and risk factors for low ovarian reserve are explored in detail in Chapter One of this thesis. Increasing age is one of the most common contributors to a woman having low ovarian reserve (203). Many couples are delaying having children, with the mean age of child bearing rising in the western countries (204). Given this trend, the number of women experiencing fertility issues due to diminished ovarian reserve is likely to rise. Studies show that women's awareness of age related fertility decline is low (205,206). In addition many women who were aware of the age related decline were under the impression that IVF would be able to reverse this entirely (207). One study showed that less than quarter of women had realistic perceptions of the chances of a successful IVF cycle (208).

There is a dearth of published studies exploring women's views about diminished ovarian reserve and ovarian reserve testing. One study showed that there is an appetite for ovarian reserve testing amongst women (209). There are also studies indicating that greater awareness of ovarian reserve will change family planning and reproductive behaviours amongst the population, despite an acknowledgement that this might not always be within every person's control (206,210,211). In order for health providers to tailor the care provided and meet the needs of women with low ovarian reserve, a greater understanding of their views and motivating factors is essential.

There is also little published evidence on women with low ovarian reserve and their acceptability of research, clinical trials and new treatments. There are some studies indicating that women of childbearing age are motivated to take part in research (212,213). However, the generalisability of these results to women seeking fertility treatment and especially women with diminished ovarian reserve is unclear. To

be able to successfully carry out research to improve outcomes for women with diminished ovarian reserve, researchers need to be able to design a research protocol which would be acceptable and appealing to this target population. In order to achieve that, a clear understanding of what drives these women to take part in research and what would be within the limits of their acceptability is needed. Furthermore, clinical trials that have engaged patients prior to the conception of the study, have been shown to have better recruitment, lower dropout and overall improved outcomes (214–216).

This study sets out to explore the views of women with diminished ovarian reserve and what they expect from a clinical trial aimed to improve their reproductive outcomes.

Methods

Ethical approval

A qualitative study was designed to explore the ideas and expectations of women with diminished ovarian reserve. Ethical approval to conduct this study was granted as part of the original ethical approval application by the North of Scotland Research Ethics Service (16/NS/0104 IRAS 204528).

Participant Selection

Purposive sampling with maximum variation sampling was carried out. Purposive sampling is a non-random sampling method where participants are chosen because of the qualities they possess (217). Purposive sampling is known to be ideal to get rich data from a limited number of participants (218). Maximum variation sampling is a subtype of purposive sampling, and describes the process by which participants are chosen because of their heterogeneity to ensure the widest variation is captured (219). This technique of sampling allows important shared patterns to emerge out of the heterogeneity and enables the capturing of both similarities and differences amongst participants (220).

Women attending a secondary care Fertility clinic at an NHS hospital (Birmingham Women's and Children's NHS Trust) were recruited. Participants were purposively chosen from women with a risk factor for diminished ovarian reserve. The risk factors for diminished ovarian reserve included participant age of 35 or above, serum follicle stimulating hormone (FSH) levels of 8.9IU/ml or above, serum anti-

mullerian hormone (AMH) levels of 5.4pmol/L or less, or a previous history of poor response to controlled ovarian hyperstimulation. There is a detailed justification of the choice of risk factors for low ovarian reserve in Chapter Five of this thesis. A detailed description of controlled ovarian hyperstimulation and poor response can be found in Chapter One of this thesis. Participants were chosen in such a way to ensure a wide range of ages, ethnicities, religions and socioeconomic factors were included. Verbal and written informed consent was obtained from each participant as part of a wider study and the full description can be read in Chapter Five.

Sampling was continued until saturation was reached and no further themes were identified (221). In total, 21 women were included.

Interview structure

Interviews were carried out at Birmingham Women and Children's Hospital. The decision to carry out interviews at this setting was made to minimise the inconvenience for participants and to avoid intrusion. All interviews were carried out by Dr Bala Karunakaran MBBS, BSc, DPMSA, MRCOG (male specialist registrar and PhD candidate). The interviewer had undergone prior training in qualitative methodology (having undertaken a course by Nuffield department of primary care and health sciences, University of Oxford).

Semi-structured interviews were used to collect data. This method was chosen as it allowed the utilisation of the investigator's a priori knowledge of women with low ovarian reserve as well as allowing flexibility to explore themes pertinent to the individual participant in detail. Semi-structured interview is considered the optimal method of interviewing when only a single sitting is possible to obtain data (222).An interview guide was produced prior to the study (attached in appendix 3). A mixture of open, probing and closed questions was used and participants were given an opportunity to express anything they considered important at the end of the interview. The interviews were conducted with either just the interviewer and the participant or if the participant elected to, the participant's partner present. Interviews were recorded and stored on a password protected computer in a secure locked room. Interviews were transcribed verbatim at a later stage. Transcripts were anonymised and analysed using NVivo 11

organisational software (NVivo qualitative data analysis Software; QSR International Pty Ltd. Version 11).

Analysis

Thematic analysis was chosen as it is considered a suitable approach which bridges the gap between qualitative and quantitative methodologies (223). A framework analytical approach was adopted. Framework analysis was first described by Ritchie and Spencer, and is a method by which clear steps are followed to produce a structured output of summarised data (224). Framework approach is used extensively in analysing semi-structured interviews (225), and has been validated in healthcare research (226).

The seven steps of thematic analysis as described by Gale was followed and an analytical matrix developed (224). The data was coded to identify overarching themes, following the principles outlined by Braun and Clarke (227). Some a priori labels were deductively identified prior to analysis, based on the investigator's prior clinical experience (228). Subsequent themes were captured inductively, using open coding, a process described by Struss, where new labels are assigned by the investigator (229). Similar labels were grouped together into themes. First ten percent of all transcripts were analysed and coded independently by a second researcher (James Cheshire, PhD student) to increase reliability and to compare and validate initially identified themes (230). An analytical framework was developed and applied to transcripts. A framework matrix was created on NVivo to chart the data. The framework matrix was analysed to interpret the underlying themes. The resultant themes and the relationship between the themes were noted by the investigators with quotes selected to best illustrate each theme and subtheme.

Results

We followed the Consolidated criteria for reporting qualitative studies (COREQ): 32-item checklist in reporting our methodology and findings (231). This is presented in Table 17.

Table 17: Consolidated criteria for reporting qualitative studies (COREQ). Modified from Tong et al., (231),.

No. Item	Guide questions/description	Reported section
Domain 1: Research team and reflexivity		
Personal Characteristics		
1. Interviewer/facilitator	Which author/s conducted the interview or focus group?	Methods
2. Credentials	What were the researcher's credentials?	Methods
3. Occupation	What was their occupation at the time of the study?	Methods
4. Gender	Was the researcher male or female?	Methods
5. Experience and training	What experience or training did the researcher have?	Methods
Relationship with participants		
6. Relationship established	Was a relationship established prior to study commencement?	No
7. Participant knowledge of the interviewer	What did the participants know about the researcher? e.g. personal goals, reasons for doing the research	Methods
8. Interviewer characteristics	What characteristics were reported about the inter viewer/facilitator? e.g. Bias, assumptions, reasons and interests in the	Methods

	research topic	
Domain 2: study design		
Theoretical framework		
9.Methodological orientation and Theory	What methodological orientation was stated to underpin the study?	Methods
Participant selection		
10. Sampling	How were participants selected?	Methods
11. Method of approach	How were participants approached?	Methods
12. Sample size	How many participants were in the study?	Results
13. Non-participation	How many people refused to participate or dropped out? Reasons?	Results
Setting		
14. Setting of data collection	Where was the data collected?	Methods
15.Presence of non-participants	Was anyone else present besides the participants and researchers?	Methods
16. Description of sample	What are the important characteristics of the sample? e.g. demographic data, date	Results
Data collection		
17. Interview guide	Were questions, prompts, guides provided by the authors? Was it pilot tested?	Appendix
18. Repeat interviews	Were repeat inter views carried out? If yes, how many?	No
19. Audio/visual recording	Did the research use audio or visual recording to collect the data?	Methods
20. Field notes	Were field notes made during and/or after the inter view or focus group?	Methods
21. Duration	What was the duration of the inter views or focus group?	Methods

22. Data saturation	Was data saturation discussed?	Methods
23. Transcripts returned	Were transcripts returned to participants for comment and/or correction?	No
Domain 3: analysis and findings		
Data analysis		
24. Number of data coders	How many data coders coded the data?	Methods
25. Description of the coding tree	Did authors provide a description of the coding tree?	N/A
26. Derivation of themes	Were themes identified in advance or derived from the data?	Methods
27. Software	What software, if applicable, was used to manage the data?	NVivo
28. Participant checking	Did participants provide feedback on the findings?	Discussion
Reporting		
29. Quotations presented	Were participant quotations presented to illustrate the themes/findings? Was each quotation identified? e.g. participant number	Results
30. Data and findings consistent	Was there consistency between the data presented and the findings?	Discussion
31. Clarity of major themes	Were major themes clearly presented in the findings?	Results
32. Clarity of minor themes	Is there a description of diverse cases or discussion of minor themes?	Results

In total 30 women were approached and invited to be interviewed. Of these, 21 women agreed to be interviewed. (Five stated that they did not have time, three were unwilling to be recorded and one gave no reason). Baseline characteristics of participants are listed in Table 18.

Table 18: Baseline characteristics of interview participants

Average Age (SD)	38.1(2.0)
Ethnicity	
White(%)	10(47.6)
Black(%)	2(9.5)
Chinese(%)	3(14.3)
South Asian(%)	5(23.8)
Other(%)	1(4.8)
Fertility status	
Primary infertility(%)	15(71.4)
Secondary infertility(%)	6(28.6)
Previous IVF(%)	2(9.5)
No Previous IVF(%)	19(90.5)
Ovarian reserve status	
Age>35(%)	16(76.2)
Low AFC(%)	8(38.1)
Low AMH(%)	9(42.9)
High FSH(%)	15(71.4)

Results are presented in four main themes:

1. Women's perception of ovarian reserve and their reasons for delaying starting a family.
2. Need for further awareness of ovarian reserve.
3. Acceptability of an antral follicle count based IVF treatment protocol.
4. Women's attitude to research and acceptability of delaying treatment to identify the optimal cycle.

Perception of ovarian reserve and reasons for delaying fertility

Most women interviewed had a good understanding of ovarian reserve and how it related to age. One participant stated:

> *"My understanding is that it deteriorates after 35 but it's not necessarily like falling off a cliff".* [Participant 06]

However, many women reported that their lack of awareness of ovarian reserve and age-related decline of fertility in the earlier part of their life led to a delay in starting a family. Many felt that fertility is a taboo subject which is not talked about in social circles.

> *"I certainly didn't hear that in my teens or in my 20s and as such I made a conscious decision to wait until 30"* [Participant 04]

> *"I knew about the ageing process but I don't think I realised it was as young as you are 35, 40. I don't think people talk about it do they"* [Participant 20]

Another participant reported that even when family members and friends want to talk about fertility, it is often considered socially unacceptable to do so.

> *"I think they did the honourable thing and not mention it to me"*. [Participant 16]

There was a trend noted among women from certain cultural and religious backgrounds reporting that it was their social norm not to discuss fertility at school or at home. One woman who grew up in Nigeria reported that discussing fertility was unacceptable when she was growing up. Another woman from an Irish background stated:

> *"I went to a catholic school and the teaching was very much about how not have sex or get pregnant (laughs)"*.
>
> [Participant 15]

The second most common reason reported for delaying starting a family was not having met the right person. These group of women generally reported that though they had regrets that they had not met the right partner earlier in their lives, they would prefer to wait for the right person instead of "*getting pregnant with anybody because my clock is ticking*" [Participant 03].

Women also reported having different priorities earlier in their life and that leading to delay of fertility. "*Priority was to study and enjoy life and travel the world*" [Participant 17]. Career also had a big bearing on the decision to delay treatment for multiple women who were interviewed.

"I chose not to have children because it was so damaging to career"

[Participant 04]

Women who claimed to work in a male dominated profession believed that once the hierarchy becomes aware that she is trying to conceive, barriers to progression will arrive.

"Opportunities would most certainly have been closed to me had I disclosed pregnancy and actually, I had to hide a number of miscarriages because of the environment, I knew the promotions wouldn't then come because it would be, "Oh, you've been trying," and the investment would stop."

[Participant 18]

Need for awareness

Most participants felt that there was a need for greater awareness about the decline in fertility with age. Some felt strongly that there should be a national education campaign. The majority felt that this education should begin at secondary school to enable women to factor this in their life aspirations.

"I perhaps would have rather been told as a teenager or something to consider my fertility in line with my biological clock". [Participant 08]

The majority of women also considered that GPs should also have an active role to play. Many felt that during their regular consultations or contraceptive checks, GPs could have initiated discussions. However, when this idea was presented to other women, some felt that this would be inappropriate.

"It's assuming that someone wants a family. Not everyone does". [Participants 17].

Acceptability of an antral follicle count based IVF treatment

Women were asked about the concept of having additional scans and blood tests to aid clinicians to start their treatment on their optimal month. In general, participants were very enthusiastic about the concept. The idea of having multiple tests to gather information rather than relying on one test intuitively appealed

to many women. Many described the thought of "not just one number, one test" [Participant 01] as being "holistically better" [Participant 21].

The participants seemed not to consider the additional burden of investigations, including additional transvaginal ultrasound scans, bothersome. Even women who had considerable distance to travel to the hospital were prepared to face the additional visits.

> *"It obviously is an inconvenience because you're not on my doorstep, and obviously I've got a child so we've got childcare to sort out, but that's fine. As I say, we've got supportive people around us. I would happily do it"* [Participant 05].

The main concern about the treatment expressed was about the potential delay to starting treatment, with concerns about the time related decline in their fertility.

> "*the more you wait longer, you get older as well, the chances are getting lower and lower*." [Participant 11]

Two of the participants categorically stated that they would not accept any form of delay in starting their IVF treatment. Most women accepted that there might be a trade-off between delay in treatment and potential better outcomes.

> "*If it made me wait a little bit longer to get hopefully a better result, then yes, or a bit more of a high percentage of it working, then definitely*". [Participant 13]

When pressed about the maximum delay in treatment they would be prepared to accept, most stated that this would be between four to six months. They emphasised on the distress caused by uncertainty and wanted a clear plan as to how many cycles they might be scanned for. One woman stated, "[I] *wouldn't mind being under treatment knowing that I had an end date*" [Participant 20].

Attitude to research

The cohort of women interviewed were very supportive of research participation. Most of them were positive about altruistic fertility research, even if it will not directly benefit them.

"*If I can help someone that, you know, might not have any children at all ... I know how that feels and I've got one ... then yeah, I think it's acceptable to do that*". [Participant 07]

Many reported that it was heartening to know that there was a high volume of research being carried out in reproductive medicine. Participants showed a great understanding of the need for research to evaluate and improve treatment options.

"*I absolutely would because obviously you want to move towards a situation where people don't continue to have treatments that are less effective*" [Participant 07].

When asked about taking part in a clinical trial about an antral follicle count based IVF protocol, most women were positive about it. One of the participants expressed her concern about being one of the first participants.

"*Err, of course you don't want to be the first one*". [Participant 20]

Others expressed concerns about the design of a randomised controlled study, with a particular dislike of the chances of ending up in the control arm of the trial.

"*The problem you've got with that is if you've got people that know each other. They're both going for the same thing, and then someone says, 'Well, I've been selected for this. You should speak to the hospital. They might be able to do you'*". [Participant 09]

The lack of choice and the uncertainty about which treatment arm one might end up in was concerning for some. However, those with concerns stated that this alone will not deter them partaking in research. Some of the participants felt that if there were additional benefits of participating in the study, such as an additional free or subsidised IVF cycle, that might motivate them to participate in the clinical trial.

Discussion

Multiple conclusions can be drawn from this study. Firstly, there are several reasons why women delay starting a family, perhaps the most significant is the lack of patient awareness of their ovarian reserve and potential for age-related decline. This lack of awareness was more prevalent in certain cultural and social groupings. Women with low ovarian reserve wished that there had been more education on this topic at an earlier stage in their life, either at school or opportunistically by GPs, although a small minority felt that such intervention was inappropriate. There is a strong desire to contribute to research amongst women with low ovarian reserve. With regards to their participation in further research, women were happy to accept a short delay in treatment if it could result in better IVF outcomes. Whilst the concept of a randomised controlled trial was not intuitively appealing, women were willing to participate in such trials if there was an overall compelling case made.

This study has multiple strengths. This study followed a validated qualitative research methodology and with its purposive sampling with maximum variation design, it successfully captured a wide range of patient voices. Sampling was continued until saturation of data was achieved. The researcher was not part of the routine care provider for patients, eliminating potential coercion. All interviews were recorded and transcribed, with 10% of the interviews being coded by two researchers independently, ensuring reproducibility.

The main weakness of this research was that it was set at a single centre. Therefore, even with maximum variation sampling and Birmingham being one of the most ethnically diverse centres in the UK, the generalisability is limited. The high acceptability of research amongst those interviewed should be read caution, as by default women who were willing to participate in an interview are likely to be positive about taking part in other forms of research. The interviews were also conducted by a researcher who was involved in wider research into antral follicle count and improving outcomes. This was deemed necessary to be able to give the participants sufficient explanation so that an informed discussion could take place. However, inherent bias of the interviewer could not be entirely eliminated.

The venue of the interview and the role of the interviewer are known to impact the responses in qualitative interviews(232,233). The interviewer was also a doctor specialising in fertility. Even though he was not part of the team providing routine care for these women, the influence of his role on the answers provided cannot be eliminated. Interviews were conducted in hospital setting rather than at participants' homes. This environment may not have put women at ease and may have influenced the flow of interviews. Results were not fed back to participants post-analysis so as to reduce patient burden. This meant that refining of themes by participants was not pursued.

Women reported poor awareness of age-related decline in fertility, especially when they were in their 20s and early 30s. These findings are similar to the findings of other studies based on the wider population (205,206). The reasons that emerged from this study as to why women are delaying starting a family, such as changes in life priorities, focus on career and a family-unfriendly workforce environment, is comparable to the findings of Mills et al (234). The general high level motivation amongst the participants to take part in research for altruistic reasons support finding of Newington et al., who interviewed patients in other medical specialities (235). The concern expressed by many of the participants towards randomisation also reflects the findings of other studies such as Jenkins et al (236). However Jenkins et al found that once additional information about the trial and the fact that participation is entirely voluntary and could be withdrawn at any time is explained to participants, the majority were willing to partake in a randomised controlled trial.(236). This also resonates with the participants of our study who stated that they would be happy to participate in randomised controlled trials despite their initial reservations, if they agree with the overall rationale for the study.

This study has identified patients' lack of awareness of their ovarian reserve and its decline. Further strategies are needed to improve the education and reduce the number of women requiring treatment with low ovarian reserve. Further research by primary care researchers to determine the best time for GPs to engage with women in discussion about fertility and ovarian reserve could be of benefit. There are multiple barriers to starting family early, including workplace discrimination and the gender pay gap, and research into strategies to overcome them are needed.

This study shows a positive attitude of patients towards being involved in research for women with low ovarian reserve. Any study designed to improve outcomes for this population should be mindful of the time pressure these women are under. When designing studies, researchers would be prudent to look at the patient's entire treatment pathway, and identify how it could be streamlined, in order to minimise any delay to starting treatment. Given the high levels of concern expressed about randomised controlled trial design, if this is chosen, then significant resources need to be allocated to explain the rationale behind the choice of design to patients to ensure high levels of participation.

Conclusion

There is a reported lack of awareness of ovarian reserve and its decline. Targeted strategies to improve education will be welcome amongst women. There is a great appetite for participating in research in women with low ovarian reserve.

Chapter seven: clinicians' perspective of low ovarian reserve and acceptability of research- a qualitative study

Preamable to chapter seven

So far in my thesis, I have shown that women with low ovarian reserve have poor reproductive outcomes. I have shown that ovarian reserve tests showed association with reproductive outcomes such as pregnancy loss and live birth rate. I have also demonstrated that AFC positively correlates with live birth and negatively correlates with pregnancy loss. A systematic review conducted as part of this thesis showed that there was menstrual cycle to menstrual cycle variation in ovarian reserve tests. The prospective cohort study I carried out showed that there is menstrual cycle to menstrual cycle variation in ovarian reserve tests in women with low ovarian reserve.

The next step in this research journey would be to test the hypothesis that identifying a month with a high AFC in a woman with low ovarian reserve, and starting treatment in that particular month, could lead to improved reproductive outcome.

In chapter six, I shared the results of a qualitative study I carried out on women with low ovarian reserve, which provided further insight into their views, and their desire to take part in clinical research. It also provided further insights into factors researchers should take into consideration when designing a clinical trial.

A successful clinical trial also needs buy in from clinicians. In this chapter, I share the results of a qualitative study I carried out exploring the views of clinicians involved in treating women with low ovarian reserve. I aim to identify their views on taking part in research. I also aim to gain further insights and experiential data on treating women with low ovarian reserve.

Contributions

Dr Bala Karunakaran- Conceived the idea, carried out interviews, transcribed, carried out analysis and wrote this manuscript.

Dr James Cheshire was the second coder and helped with analysis.

Prof Coomarasamy proof-read the manuscript and provided substantial edits.

Abstract

Study Question

What are the views of clinicians who treat women with low ovarian reserve with regards to the current treatment options that are available, the current gaps in care and their appetite to support research?

How acceptable is a research trial based on serial ultrasound scanning of women with low ovarian reserve to identify the optimal month to start IVF treatment?

Summary Answer

Clinicians find consultations with women with low ovarian reserve challenging. There is a keen interest to contribute to research amongst IVF clinicians. Most clinicians find a research study based on serial ultrasound scans to identify the optimal month to start IVF treatment acceptable.

What is known already?

There is a high level of anxiety amongst IVF providers, especially related to poor outcomes in patient. There is research showing collaborative priority setting with the involvement of all stakeholders lead to better clinical outcomes.

Study size, design, and duration

Qualitative study over a period of three months.

Participants/materials, setting, methods

Purposive sampling with maximum variation amongst experts. 10 doctors, six embryologists and three nurses from locations across England. Semi-structured interviews recorded and transcribed verbatim. Thematic analysis carried out using the framework methodology.

Main results and the role of chance

Clinicians found these consultations particularly challenging because of the perceived high expectations of the women. Unexpectedly low ovarian reserve test results often cause distress to women and clinicians. There is a lack of acceptable alternative treatment options, as women often find using donor eggs unpalatable. There are cultural and religious factors which affect the acceptability of donor eggs.

There is strong support for facilitating research amongst reproductive clinicians. The ideal research project designed to with a clinically relevant research question, simple in its design, not be resource intensive, be multi-centred and have the potential to make changes to clinical practice. It was felt that clinicians with an established prior relationship with the participant were better placed to recruit to studies.

Clinicians found the concept of a research study based on scanning for consecutive months and identifying the optimum cycle for women with low ovarian reserve appealing and most were positive about wanting to be part of such study if offered. Practical issues for consideration included additional psychological, physical and logistical burden posed by delaying the onset of treatment and need for the availability of sonographers with the expertise.

Limitations, reasons for caution

Study carried out in England only so generalisability is limited

Wider implications of the findings

Fertility care providers should ensure that there is provision for longer consultations supported by staff trained in counselling to better support women with low ovarian reserve. There is a need for a public health campaign to reduce the stigma of donor eggs, especially amongst some ethnic and religious communities. This chapter also sets out a criteria which is useful to have in consideration when designing research projects.

Introduction

There are very few qualitative studies that have been published in the field of reproductive medicine that study the perspectives of clinicians. This is despite the fact that there is evidence showing that in the IVF setting, when the needs of the healthcare staff are taken into consideration, patient outcome improves (237). There is evidence that anxiety levels are higher in doctors practising reproductive medicine in comparison to other doctors and that unsuccessful outcomes of IVF treatment is one of the main stressors (238). It has been shown that women with low ovarian reserve have lower chances of success with IVF treatment (44,239,240). Therefore it is likely that treating women with low ovarian reserve can be technically as well as emotionally challenging to clinicians. Research into clinicians treating women with low ovarian reserve has focused on establishing clinical consensus on definition and treatment options, and has failed to delve deeper into understanding their experiences and views (34,241).

There is a need for further research to improve reproductive outcomes for women with low ovarian reserve. Boaz et al., argued that involvement of stakeholders in research is crucial in bridging the gap between research production and the research usage, resulting in greater impact (242,243). There is evidence that involving healthcare professionals in the design, conduct and interpretation of clinical trials yields better clinical outcomes (244,245). Esmail et al., make the case that stakeholder involvement should be included in setting the priorities for research and should be embedded into all research projects (246). Cook et al., have made the case that collaborative priority setting is applicable for research conducted within the NHS (247).

This study aims to investigate healthcare providers' understanding of women with low ovarian reserve, and what they perceive their patients expect from treatment and research. This study will examine/assess clinicians' attitudes towards collaborating in fertility research and the feasibility of conducting a research to improve reproductive outcomes in women with low ovarian reserve.

Methods

Participant Selection

Purposive sampling of experts was carried out. Purposive sampling is a non-random sampling method where participants are chosen because of the qualities they possess (217).Purposive sampling is known to be ideal to get rich data from limited number of participants (218). Expert sampling is carried out when a group of participants are approached because of their expertise in a particular field (217,248). Within the group of experts, maximum variation sampling was used. Maximum variation sampling is a subtype of purposive sampling, and describes the process by which participants are chosen because of their heterogeneity to ensure the widest variation is captured (219). This technique of sampling allows important shared patterns to emerge out of the heterogeneity and enables the capturing of both similarities and differences amongst participants (220).

Doctors, fertility nurses and embryologists practising in both state funded institutions and private sector were approached for interview. Participants were also selected amongst attendees at fertility conferences as well as by contacting IVF units. Sampling was continued until saturation was reached and no further themes were identified (221).

Interview structure

All interviews were carried out by Dr Bala Karunakaran (specialist registrar and PhD candidate, Male, Qualifications MBBS, BSc, DPMSA, MRCOG). The interviewer had undergone prior training in qualitative methodology (Attended course by Nuffield department of primary care and health sciences, University of Oxford).

Semi structured interviews were used to collect data. This method was chosen as it allowed the utilisation of the investigator's a priori knowledge of women with low ovarian reserve as well as allowing flexibility to explore themes pertinent to the individual participant in detail. Semi structured interview is considered the optimal method of interviewing when only a single sitting is possible to obtain data (154). An interview guide was produced prior to the interviews. Mixtures of open, probing and closed questions

were used and participants were given opportunity to express anything they considered important at the end of the interview. The interviews were conducted at a location of convenience for the participant. Interviews were recorded and stored in a password protected computer in a secure locked room. Interviews were transcribed verbatim at a later stage. Transcripts were anonymised and analysed using NVivo 11 organisational software.

Analysis

Thematic analysis with framework approach was chosen, as this a validated qualitative research method which has been used in many settings to analyse semi-structured interviews, including in the healthcare setting (223,224). The justification for using this method of analysis in qualitative interviews is outlined in Chapter Six of this thesis.

We followed an identical methodology as described in chapter six of this thesis.

Results

We followed the Consolidated criteria for reporting qualitative studies (COREQ): 32-item checklist in reporting our methodology and findings(231). This is presented in Table 19.

Table 19 Consolidated criteria for reporting qualitative studies (COREQ). Modified from Tong et al., (199)

No. Item	Guide questions/description	Reported section
Domain 1: Research team and reflexivity		
Personal Characteristics		
1. Inter viewer/facilitator	Which author/s conducted the interview or focus group?	Methods
2. Credentials	What were the researcher's credentials? MBBS, BSc, DPMSA, MRCOG	Methods

3. Occupation	What was their occupation at the time of the study?	Methods
4. Gender	Was the researcher male or female?	Methods
5. Experience and training	What experience or training did the researcher have?	Methods
Relationship with participants		
6. Relationship established	Was a relationship established prior to study commencement?	No
7. Participant knowledge of the interviewer	What did the participants know about the researcher? e.g. personal goals, reasons for doing the research	Methods
8. Interviewer characteristics	What characteristics were reported about the inter viewer/facilitator? e.g. Bias, assumptions, reasons and interests in the research topic	Methods
Domain 2: study design		
Theoretical framework		
9.Methodological orientation and Theory	What methodological orientation was stated to underpin the study?	Methods
Participant selection		
10. Sampling	How were participants selected?	Methods
11. Method of approach	How were participants approached?	Methods
12. Sample size	How many participants were in the study?	Results
13. Non-participation	How many people refused to participate or dropped out? Reasons?	Results
Setting		
14. Setting of data collection	Where was the data collected?	Methods

15.Presence of non-participants	Was anyone else present besides the participants and researchers?	Methods
16. Description of sample	What are the important characteristics of the sample? e.g. demographic data, date	Results
Data collection		
17. Interview guide	Were questions, prompts, guides provided by the authors? Was it pilot tested?	Appendix
18. Repeat interviews	Were repeat inter views carried out? If yes, how many?	No
19. Audio/visual recording	Did the research use audio or visual recording to collect the data?	Methods
20. Field notes	Were field notes made during and/or after the interview or focus group?	Methods
21. Duration	What was the duration of the inter views or focus group?	Methods
22. Data saturation	Was data saturation discussed?	Methods
23. Transcripts returned	Were transcripts returned to participants for comment and/or correction?	No
Domain 3: analysis and findings		
Data analysis		
24. Number of data coders	How many data coders coded the data?	Methods
25. Description of the coding tree	Did authors provide a description of the coding tree?	N/A
26. Derivation of themes	Were themes identified in advance or derived from the data?	Methods
27. Software	What software, if applicable, was used to manage the data?	NVivo

28. Participant checking	Did participants provide feedback on the findings?	Discussion
Reporting		
29. Quotations presented	Were participant quotations presented to illustrate the themes/findings? Was each quotation identified? e.g. participant number	Results
30. Data and findings consistent	Was there consistency between the data presented and the findings?	Discussion
31. Clarity of major themes	Were major themes clearly presented in the findings?	Results
32. Clarity of minor themes	Is there a description of diverse cases or discussion of minor themes?	Results

In total, 25 clinicians were approached for interview, of which 19 agreed to be interviewed. The six who declined to be interviewed gave their current busy schedule as the reason. The baseline characteristics of the interviewees is given in the table below.

Table 20: Baseline characteristics

Profession	**Number**	**Percentage**
Doctor	10	47.4%
Nurse	6	31.6%
Embryologist	3	15.8%
Sector		
Private	2	10.5%
State	12	63.2%
Both	3	15.8%
Gender		

Male	12	63.2%
Female	7	36.8%
Ethnicity		
White	10	52.6%
Black	3	15,8%
Asian	4	21%
Other	0	0%

We present the results of our qualitative study under three headings: Treating women with low ovarian reserve, qualities of an ideal research study and the acceptability of an antral follicle count based research study.

Treating women with low ovarian reserve

The consultations

Clinicians universally felt that consultations with women with low ovarian reserve were difficult and often provoked anxiety about how the patient might receive the news and how best to deliver the consultation with least amount of distress caused to the patient as possible.

> "*Erm, I'm anxious in terms of how the patient's going to receive the news of low ovarian reserve. Erm, and I'm also thinking about, you know, what sort of success rate should I quote either with natural conception or with IVF*." (Doctor 03).

Many doctors compared their consultations with women with low ovarian reserve to consultations they previously had with patients who were being informed of a newly diagnosed cancer.

> *"It's a difficult ... first of all it's a difficult consultation if without warning you are telling a young patient your reserves are low, ... it's like breaking the bad news of cancer to young women"* (Doctor 05).

Many clinicians, particularly those who worked in government funded IVF clinics, identified multiple aspects of these consultations as particularly troublesome. Firstly, by the time ovarian reserve tests are done, a patient is often committed in their fertility journey. This is because AMH is self-funded by patients and therefore often only done just prior to embarking on IVF treatment. AFC is not routinely done and often only performed at the beginning of fertility treatment. Many clinicians reported that women found the news of low ovarian reserve and its implications difficult to receive at this stage of their treatment.

> *"Sadly, sometimes by the time they get to us it's not been mentioned that they've got a low, erm, a low ovarian reserve and it can be quite a shock when they get to the fertility services and find out actually you're gonna have to be on the top dose of medication because you've got low ovarian reserve, the likelihood of you responding well is quite slim, you know and, and they're hearing this information for the first time which is not ideal."* (IVF Nurse 03).

> "*You know, it's very challenging because patients have a very high set of expectations when*
>
> *they come to see anyone in the fertility service. And even when we find out that they have got low ovarian reserve those expectations still stay quite high. So, I always find it quite a challenging consultation in general"* (Doctor 09).

It was also felt that younger women struggled with the notion of having low ovarian reserve in comparison to older women. A senior fertility nurse also remarked that people's perception of what is considered as older in terms of fertility has been shifting, with increasing expectations, making consultations about low ovarian reserve harder.

> *"Probably, because they weren't expecting it, but then I think age, it's got a little bit different as I've worked here for years, that people now who are 39, 40 don't consider themselves old and we know that for fertility they are. So, I think even for*

them, when it's lower than what it should be for their age they're still a bit shocked and for them, they know their chance ... I think having a very poor AMH at 40, you probably don't do as well as a little bit lower younger." (IVF Nurse 09).

Many also commented on cultural differences in how the diagnosis of low ovarian reserve was perceived, with people from some ethnic backgrounds placing great emphasis on being fertile.

"Asian culture tends to, erm, be a lot of denial, it's not, can't be a problem, because you hit the devastation then of, erm, the, the wider impact because it doesn't only affect their fertility, it can affect their marriage, their family relationships; it's a very wide reaching thing" (IVF Nurse 04).

Egg Donation as an alternative

IVF with donated eggs is a potential option to help women with low ovarian reserve conceive. However many couples do not find this acceptable.

"The women's expectations are to go through IVF treatment even if they have diminished ovarian reserve and to take home a baby that they are the genetic mother and father of." (Doctor 04)

According to clinicians, couples found the option of egg donation more acceptable only after going through multiple cycles of IVF attempting to retrieve and use their own eggs, even if they knew it was futile from the outset.

"For our other patients having erm, the, the concept of egg donation, if they've been in treatment several cycles and it's gradually this concept of egg donation is, is brought to their attention, they gradually accept it because they know that they're not getting anywhere with their own eggs and that's been introduced to them at an early time and, erm, gradually over time, they become accustomed to the fact that this is their way forward." (IVF nurse 09)

Cultural differences were again noted, with many of the clinicians experiencing that women from certain ethnicities and religious grouping finding egg donation unacceptable.

> *"So, to give an example a lot of Asian Muslim patients tend not to even consider donor egg IVF, but I've always been surprised by a small number of patients who would."* (Doctor 05)

Even for ethnic minority women for whom donor eggs are acceptable, they were often specific about the characteristics of the donor they would find acceptable, and in general finding a suitable donor for an ethnic minority often proved to be more challenging.

> *"there is still the I don't want a donor from here, from here, from here, this type of donor, that type of donor and religion as well comes into it. So it can be very, very difficult, there's very long waiting lists and particularly as now going abroad is not as easy"* (IVF Nurse 01)

Counselling and support

Many of the interviewees felt that their local NHS hospital clinics were not designed to deliver the news of low ovarian reserve to women and they did not receive adequate support and counselling after the news was delivered.

> *"I think the doctor's appointments where they discuss results with them are too short, they're too rushed, when they tell them the bad news you know they got, maybe they don't go into it in enough detail because they haven't got the time"* (Nurse 02)

Qualities of an ideal research study

Participants seem to value research as important. Participants placed high emphasis on research that led to directly answering a clinically relevant question.

"Honestly, I like research that I think, "Gosh, I wish I'd thought of that". So, something that's novel, something that will change the way we practice. And, generally something that's relatively simple to do so it comes with a nice clear outcome. So sometimes you go to conferences and someone presents you with an idea and you think, "Oh my God, that would be great. If we only knew the answer to that then it would change what I would do for every patient I see". That's the sort of research that excites me. Yeah" (Doctor 03).

IVF nurses and embryologists placed considerable emphasis on being able to understand rationale behind research before they can sign up to it. They reported that they would struggle to support and recruit if they do not understand the study fully or are able to explain it.

"When you're not entirely sure of the benefits of the trial, it's like trying to sell something you're not, you don't know, you know, here, here's a glass, I can't tell you whether it's made of plastic or glass or china" (Embryologist 02).

"*It's very difficult to encourage people to, to join a trial when it's not easy to talk to them about it*" (IVF Nurse 05).

Clinicians also placed significant emphasis on the opinion of their peers and trusting clinical trials that have been talked about in their professional circles.

"*Erm, if the trial looked as though it, it's, if, if it's say multi centre trial and I have a lot of, of contacts outside of this unit and other fertility centres and obviously we talk erm, so trial, trials that are multicentred tend to be the ones that are, are more, erm, widely talked about because there's lots of people talking about them which makes it easier to discuss it with the patients because you get more, erm, more reviews.*" (IVF Nurse 07).

Participants also expressed preference for research studies that were simple in design and not onerous with regards to additional burdens such as follow up after the trial that it places on clinicians.

"Simplicity of the treatment, or the intervention certainly comes into play, so the more complex the intervention and follow up the less likely I think trials would be, would be taken up" (Doctor 05)

There was also the view that research studies that have clinicians recruiting, rather than dedicated research staff such as a research nurse, was preferable because the woman has a prior relationship with the clinician and trusts the information from them. However there was a recognition that once the initial recruitment had taken place, the support was needed from research staff to do the necessary paperwork and free up clinicians' time.

"*I think it's better if you already know the patient. I think so. Having been in both situations myself, I get the impression that patients are more likely to trust what you say and therefore are more likely to invest time in what you're offering to them. They almost have an investment in you as much as you have an investment in them. Whereas, if you are just coldly approaching someone, I think it's very difficult sometimes to get them to take in and trust you straight away.*" (Doctor 04).

Many participants felt that whilst women are motivated by altruism to participate in research, giving them something as an extra incentive often helped to offset the additional burden placed by research.

"If there are any incentives in there for them, erm, to, to take part in a trial I think they'd bite your hand off, erm, but a lot of the trials there are no incentives and so people don't feel the obligation to, to take part in it. You know, if they were getting a free cycle of treatment by waiting and having this, that and the other or if they're having some kind of benefit on the, the drugs or what have you, I think you'd get a better uptake" (IVF Nurse 07).

The ideal fertility research study

Research question should be clinically relevant and likely to change practice.

Rationale behind the study should be understood by all clinicians.

Design should be simple and not resource intensive.

Preferably multi-centred.

Offer incentives for participants to take part

Acceptability of an antral follicle count based IVF protocol study

All clinicians, apart from one (one doctor), were positive about the concept of a study looking into identifying the optimum month to start treatment based on AFC and felt that patients would find this appealing.

> " *if you take time to explain the science, and take time to explain that you are personalising the treatment for them, I think most of these women will sign up to it. When your chances are low to start with, they want to give it the best shot*" (Doctor 07).

> *"I like the concept of it, actually. It's quite a sensible thing. So you might, I suppose, see them in the first assessment. See that they've got a low ovarian reserve, maybe only got four antral follicles or something. And then they come round to start treatment. You scan them. They've only got two antral follicles. You might be like maybe should we just wait another month and see if you've got a few more next month? And then start stimulation then."* (Doctor 02).

However the biggest limiting factor that was identified was the potential delay in starting treatment and how some women may find that unacceptable. The current long waiting lists in NHS results in women

often taking more than a year from their initial consultation to starting treatment. Introducing further potential delays when they are about to start treatment could prove troublesome.

> "*Whenever you go through a consent signing appointment, the patient feels that they want to start treatment now, if not yesterday and so having to continuously wait to be scanned in the early part of the cycle, am I gonna start, aren't I gonna start, and to be told again second month, third month, no we're not gonna go with this one, we're gonna scan you again next month t could come up with some, erm, obstacles*" (IVF Nurse 02).

Participants felt that this could be partially offset by giving women clear timelines and a plan on the maximum time they may have to wait for to identify their optimum month to start treatment.

> *"Have, having a timescale so it's like, erm, five (follicles) might be your best but if we don't see five (follicles) within four consecutive scanning cycles – we start treatment. Something like that would reassure" (IVF Nurse 02).*

One doctor felt that a small proportion of his patients with very busy life schedules, would place more importance on starting treatment when it fitted their lives rather than when it might be the optimum time to treat.

> "*Yeah. And the other thing that makes it difficult to plan is a lot of people plan IVF around their life and their career, like if they're going on holiday. A typical example is teachers, isn't it, who would rather have it over their summer holidays or whatever rather than in term time. Things like that that you might not necessarily think about when you're just thinking purely medically about what's going to give them the best response. Sometimes people time their IVF to what suits the rest of their life, so those people will be harder to recruit or to put off.*" (Doctor 08)

Others identified practical issues that might be created by the study, such as their departments currently not routinely performing AFC and therefore not having the right people with the right skills.

"*You have a variety of scanning skills. We've got new scanners that are, are learning, we've got some very experienced scanners. We're just about to change the way that we're doing our scan clinic so I, I'm not sure if you can always guarantee that you're gonna have a, erm, a skilled practitioner able to do specific scans*." (IVF Nurse 04).

Uncertainty of when a particular woman will be starting her treatment could also be problematic for departments, as often there are long waiting lists and decision on who starts treatment is made months in advance.

"*If you are talking about one or two women a month, in the trial idea you talked about, we can just about cope. Anything more than that will be difficult. At the moment we have a four month wait from signing consent to starting down reg and I can't justify it*" (Doctor 01).

The one doctor who said they did not like the concept of this study, gave their previous negative experience, and their belief that varying the way women with low ovarian reserve are treated does not change outcome, as the reason for rejecting this study.

"*Personal experience, poor responders who do not want to give the time. You want to get on with it. Secondly, whatever follicle you use, whatever dose of medication you give, it doesn't seem to make much difference. Honestly, I think it will be a little bit negative consultation to start with because I'm wasting patients' time. I am sceptical.*" (Doctor 15)

Discussion

Principle findings

Our study shows that clinicians find consultations with women with low ovarian reserve difficult. They often compare those to consultations where they had to break bad news such as conveying a cancer diagnosis. Clinicians found these consultations particularly stressful because of the perceived high

expectations of the women. Ovarian reserve tests such as AMH and AFC are often performed at a later stage of the woman's fertility treatment, which often meant unexpectedly low results needing change of treatment plans, causing distress to women. Clinicians find treating women with low ovarian reserve difficult because of the lack of acceptable alternative treatment options. Women often find the idea of using donor eggs unpalatable at the onset. However, after having then had few unsuccessful attempts to have IVF treatment with their own eggs, they become more accepting of donor eggs. Younger women and women from certain cultural and religious background tend to find the concept of donor eggs unacceptable.

Reproductive practitioners felt a strong affinity towards research. Their ideal research should be designed to centre around a clinically relevant research question, should be simple in its design, should not be resource intensive, should be multi-centred and should have the potential to make changes to clinical practice. They felt that research projects which were easier to understand and explain and offered some sort of additional incentive to the participant were the easiest to recruit to. It was universally felt that clinicians with an established prior relationship with the participant were better placed to recruit to studies in comparison to designated research staff who had never met the woman before.

Clinicians found the concept of a research study based on scanning for consecutive months and identifying the optimum cycle for women with low ovarian reserve appealing and most were positive about wanting to be part of such study if offered. However they did have some issues for consideration, including the additional burden posed by delaying the onset of treatment, need for the availability of sonographers with the requisite training to perform AFC, and the uncertainty created by not knowing exactly when these women might start their treatment and the resource implications resulting from that. It was also expressed that some women would value knowing exactly when their treatment starts more than any potential clinical benefit brought by delaying treatment, and for these women this study might not be suitable.

Comparison with existing literature

There are very few published literature exploring how clinicians find treating women with fertility issues. There are no published papers to our knowledge exploring clinician's experience of treating women with low ovarian reserve. Universally all clinicians interviewed considered their consultations with women with low ovarian reserve difficult, often due to having to break bad news and being unable to meet the high expectations of patients. This supports the findings of Boivin et al., whose study of IVF clinicians identified these two factors amongst some of the top stressors for the profession (249). The denial of the diagnosis of low ovarian reserve and the expectations of women for treatment to succeed even against the odds reported in our study, is similar to the findings of Peddie et al., who explored the decision making process of women with subfertility (250). The reluctance of acceptance of egg donation amongst some cultural backgrounds such as Muslim women has been reported in literature (251). Complex reasons have been cited for this reluctance, including interpretation of religious rulings and strong cultural emphasis on kinship and lineage (252).

Our finding of clinicians preferring research studies with a simple design over complex designs resonates with the findings of Thoma et al., who studied the research preferences of practising surgeons (253). The Cochrane review by Mapstone et al., supports our finding that there is a higher recruitment rate for studies which incentivise participants (254). However such incentives should be deployed with care to avoid ethical pitfalls such as coercion and exploitation, as cautioned in the published review by Tishler et al., (255). The cultural difference in accepting the diagnosis of subfertility and the perceived stigma surrounding gamete donation has been previously reported by Culley et al., (256).

Strengths and limitations

This study has multiple strengths. This study followed a validated qualitative research methodology and with its purposive sampling design, it was set to capture a wide range of clinician views. Clinicians were sampled from multiple locations across the UK, and from a range of different professionals such as doctors, nurses and embryologists were included, ensuring adequate representation and generalisability of findings. Sampling was continued until saturation of data was reached, increasing the likelihood that

majority of the key themes were captured. All interviews were recorded and transcribed, with 10% of the interviews being coded by two researchers independently, ensuring reproducibility.

The study is limited by the weaknesses of its chosen methodology. Whilst purposive sampling is considered best suited to reach saturation efficiently, it relies on the judgement of the investigator in selecting samples. Therefore the impact of conscious and unconscious bias on the selection of subjects cannot be excluded. Transcripts were not returned to interviewers to verify, which can pose some questions about the validity of data. All interviews were conducted by the same researcher. Thus any bias or assumptions introduced by the interviewer's questions or questioning style could not be eliminated.

Implications of findings

Young women and women from certain ethnic and religious backgrounds can find both the diagnosis of low ovarian reserve and the concept of egg donation difficult to accept. Therefore fertility centres may benefit from commissioning counsellors with expertise in catering for the needs of this specific population. Working with community leaders and religious elders to eliminate the stigma attached to fertility problems may also prove to be beneficial. Clinicians reported that ovarian reserve tests are performed after the patient and the clinician had committed to a particular treatment option, with the consequent change of treatment plans causing patients distress. Therefore it will be prudent to perform these tests early in the patient journey to better manage expectations. Our study reported perceived difficulties in the diagnosis of low ovarian reserve being conveyed to patients at routine fertility clinics, where the clinicians had considerable time constraints. Seeing these women at a dedicated appointment with the support of staff trained in counselling could alleviate some of the distress expressed. Some of the difficulties in consultations arose from perceived unrealistic expectations from women, and particular poor awareness of age related decline in fertility. Further research into finding ways to successfully convey to the wider population the success rates of IVF and about how this might be impaired by age related decline in ovarian reserve may be warranted.

Our study showed enthusiasm among IVF clinicians to participate in research. To increase engagement with clinicians, those planning research studies should make them simple in design, not very onerous and

multi-centred. Researchers should also place efforts to engage with not only the principle investigators, but all frontline clinical staff and ensure that they have a good understanding of the underlying rationale for the research project. A partnership model, with the clinician know to the patient taking the lead in recruitment, supported by dedicated research staff for administration and follow-up, is likely to enhance recruitment.

With regards to a hypothetical study identifying the optimum month to start IVF based on AFC, efforts should be made to recruit women at an early stage of their patient journey, ideally at the very first outpatient clinic appointment, to offset any delay caused by serial scanning. Clear communication of the protocol, and deploying the shortest maximum duration for which a woman might have to wait for to identify the optimum month to start her IVF treatment, would also help to allay some of the clinician and participant anxiety. A multi-centred approach would also be recommended, so that the number of women who may potentially start their IVF treatment at any given month could be restricted to small numbers, reducing resource allocation concerns borne by the uncertainty of when a women might start her IVF treatment. Researchers should also either restrict the selection of centres to ones where most sonographers and clinicians have the expertise in performing AFC, or make provisions to ensure adequate training is made available prior to commencing the study.

Clinicians also identified the role of additional incentives for participating in research and how this might encourage participants to take part. This has ethical implications. There is a plethora of published literature of the pitfalls of actual or perceived coercion in research(257,258). Care needs to be taken to ensure that research study designs are scrutinised to exclude intended or unintended coercion and research recruiters are trained to identify such coercive factors and ensure that they play no part in recruiting.

Conclusion

Clinicians find consultations treating women with low ovarian reserve difficult. The acceptability of donor egg IVF is low, particularly amongst young women and those from certain cultural and ethnic background. There is considerable enthusiasm to participate in research amongst IVF clinicians.

Participation in research could be encouraged further by ensuring the design of the study meets the recommendations made in this article.

Chapter eight: Conclusion to thesis

In chapter two of this thesis, I shared the results of a systematic review of published studies, studying the menstrual cycle to menstrual cycle variation in the ovarian reserve tests of AFC, AMH and FSH. The method of reporting inter-cycle variability in the constituent studies were heterogeneous in nature, leading to a narrative synthesis. This review demonstrated that there is an inter-cycle variation in ovarian reserve tests.

In chapter three of this thesis, I shared the results of a systematic review and meta-analysis I carried out. In this study, I studied the relationship between AFC, AMH, FSH and pregnancy loss. This study showed a statistically significant association between reduced ovarian reserve (low AFC, low AMH and high FSH) and increased chance of pregnancy loss.

In chapter four of this thesis, I shared the findings of a large cohort study I carried out, analysing the data set of a large IVF provider. I studied the relationship between AFC, and important reproductive outcomes such as live birth and pregnancy loss. This study showed a positive association between AFC and live birth rate and a negative association between AFC and pregnancy loss.

Using chapters three and four, I have demonstrated a strong association between the ovarian reserve tests and reproductive outcomes. Chapter two showed the existence of inter-cycle variation in ovarian reserve tests, which may have clinical significance. This is because, as I described in my illustrative example (chapter one), a woman with a low ovarian reserve might have an AFC of four one month and seven in another. In such woman, the first month, where there was only an AFC of four, there could be a low ovarian response, resulting in fewer eggs (for example four eggs). This would result in fewer fertilised eggs (for example two eggs), further leading to only one embryo suitable for implantation. This could be contrasted with another month in the same woman, where there was an AFC of seven and therefore where she may have a better ovarian response. This would lead to more eggs (for example seven eggs), and more fertilised eggs (for example four fertilised eggs) resulting in two or three embryos which were suitable for implantation. The clinical impact of inter-cycle variation is only likely to be of importance in women with low ovarian reserve, as an increase from one embryo to three embryos to choose from has the potential to improve a woman's reproductive outcome, whilst an increase from ten to 14 is unlikely result in a different clinical outcome.

The next logical question that therefore arises is whether there is cycle to cycle variation in ovarian reserve tests in women with low ovarian reserve.. There are no published research papers studying this subset of women. Thus, I carried out a prospective cohort study, recruiting women with at least one know risk factor for low ovarian reserve. The results of this study are conveyed in chapter five of this thesis. Over three consecutive menstrual cycles at the early follicular phase, I took blood tests for AFC and AMH and undertook transvaginal ultrasound scans for AFC.. The mean AFC Max variation, AMH Max variation and FSH Max variation were 3.87, 9.61 pmol/L and 6.48 IU/ml respectively. When analysis was restricted to women who we deemed to have low ovarian reserve (AFC of 10 or less, AMH of 5.5 pmol/L or less or FSH of 8.9 IU/ml or more for one or more measurements), the mean AFC Max variation, AMH Max variation and FSH Max variation were 2.4, 1.2 pmol/L and 3.8 IU/ml respectively. This establishes that in women with low ovarian reserve, there is variation in ovarian reserve tests between menstrual cycles. Amongst the three ovarian reserve tests, AFC emerged as the test with the most likely significant inter-cycle variation, as even when the analyses was restricted to women with proven low ovarian reserve, there was a maximum inter-cycle variation of 2.4 follicles.

The next step in the research journey to improve reproductive outcomes in women with low ovarian reserve would be to test the hypothesis that starting IVF treatment on the month with the highest AFC for a particular individual would result in better reproductive outcomes. However, before conducting such a study, one needs to assess the acceptability of such treatment and research amongst the key stakeholders.

In chapter six, I describe a qualitative study with purposive sampling (maximum variation method), which I conducted to study women with low ovarian reserve. The thematic analysis showed that women with low ovarian reserve have a strong desire to take part in research, even when it may not be of direct benefit to them. They found the concept of an antral follicle count based IVF treatment intuitively appealing and if offered would be keen to have it. They also found the concept of taking part in a research study to establish that choosing the optimum month to start IVF treatment based on AFC agreeable. However they had some reservations about taking part in a randomised controlled trial, as they felt uncomfortable with the role chance plays in determining their treatment. They also had concerns about any increase in the time it might take to receive treatment as a result of taking part in the study. This study offered further insight into women with low ovarian reserve. It highlighted their lack of awareness of age related decline in

fertility in the early part of their lives, the perceived impact of negative discrimination towards women who become pregnant at workplace has on delaying conception and the need for awareness of age related fertility decline in the general population.

In chapter seven of this thesis, I shared the results of a qualitative study with purposive expert sampling of fertility clinicians, including doctors, nurses and embryologists. The thematic analysis shows a strong desire to support and participate in clinical research. Clinicians found the concept of using serial measurements of AFC to identify the best month to start IVF treatment agreeable. They also expressed interest in recruiting for such study and felt that the population they treat would find it acceptable. They expressed preference for a study which is multi-centred, not very resource intensive and offered some sort of incentive for women to take part in. They cautioned that any delay to start of the treatment should be kept to a minimum for it to be acceptable to participants and clinicians. Practical considerations were also mentioned, such as training of people with the expertise to conduct AFC and keeping number of women who were undergoing ultrasound scans to a minimum.

This study of clinicians also provides a number of interesting observations, including that IVF clinicians found the consultations with women with low ovarian reserve difficult, due to the limited acceptable alternatives available and the low success rate of IVF in this cohort. They observed cultural differences in women, with a high prevalence of stigma attached to being infertile in certain ethnic groups and low acceptability of egg donation in some cultural and religious groups. It was also noted that some women were unaware of the scale of age related fertility decline. For some women, it was observed that there was an incorrect assessment of their fertility potential and unrealistic expectation of the ability of IVF to achieve conception. Some clinicians called for a public campaign to increase awareness.

Implications for future practice

There are a number of clinical observations that can be made from the findings of my thesis. Firstly, my work establishes cycle to cycle variation in ovarian reserve tests, in both the overall population of women with subfertility and in women with low ovarian reserve. Therefore, clinicians should exercise caution when using ovarian reserve test measurements in a single menstrual cycle to make a clinical diagnosis and decisions about whether to proceed with IVF treatment or not. Our study shows ovarian reserve tests are

often performed at a late stage of the patient journey, after both the patient and clinician have committed to proceed with treatment. At this stage, clinicians find it difficult to discuss abandoning treatment or considering alternatives such as egg donation. As such, clinicians should consider requesting these investigations as early as feasible.

My thesis also establishes the relationship between ovarian reserve tests and pregnancy loss in women having IVF treatments. As pregnancy loss is associated with significant psychological morbidity for both partners, women with low ovarian reserve should be thoroughly counselled about their higher risk of pregnancy loss prior to commencing IVF treatment. Our qualitative study also showed that many clinicians felt that they were supporting women with low ovarian reserve in suboptimal conditions. Therefore care providers should make every effort to ensure that women with low ovarian reserve can be seen at a clinic with longer appointment windows and by staff with training on counselling.

Implication for public policy

This thesis shows that there is a strong lack of awareness of age related decline in ovarian reserve amongst women. There was a sense of being let down by society amongst many of the women interviewed. Therefore it may be important for policy makers to look at addressing this, perhaps by introducing this as part of the national curriculum of education.

Cultural and religious stigma of being infertile and accepting egg donation was reported by participants. Clinicians had also reported a shortage of eggs donated by ethnic minority women. Whilst some assumptions have been made about why some cultural and religious groups may have negative attitudes to certain aspects of fertility treatments, in general this area is poorly understood. Policy makers need to carry out further qualitative research involving participants from these communities to gain further understanding in this issue. Policy makers should consider suitable strategies to engage with community leaders and identify solutions to eradicate stigma and improve the community support for women.

Gender discrimination and pressures in workplace to not conceive were reported as reasons for delaying conception. This issue needs addressing to stop women having to choose between a career and starting a family. This fits in with the wider gender pay gap research which is taking place in the UK.

In the introduction of this thesis I shared some of the evidence showing the psychological, emotional and social costs of infertility. The qualitative research I carried out also demonstrated the wddwctof infertility on women's wellbeing. The single biggest barrier for accessing fertility treatment in the UK is it's prohibitive cost. There is also variation in what is funded in the UK, with some healthcare authorities funding multiple IVF treatment cycles whilst others have stopped funding entirely. This is often described as 'postcode lottery'. Such healthcare inequalities, especially within the same country, is inherently unfair. Policy makers should engage with relevant stakeholders including political leaders to end such postcode lottery.

Implications for research

My thesis establishes the relationship between low ovarian reserve and pregnancy loss in women who have assisted conception. Further research is needed to see if such a relationship exists in women who conceive naturally.

In my thesis, I have shown that inter-cycle variation exists between ovarian reserve tests. I have shown that this variation is preserved in women with low ovarian reserve, best demonstrated with AFC. I have shown that from cohort data, incremental increases in AFC results in higher live birth rate and lower pregnancy loss rate. I have shown that an IVF treatment protocol based on serial ultrasound scans being performed to establish the optimal month to start IVF treatment, is acceptable to most clinicians and women with low ovarian reserve. I recommend a study of women who have serial AFC measurements to identify and start treatment on an optimum month, and establish whether such an intervention results in better IVF outcomes. If such treatment protocol is proven to yield better reproductive outcomes, a cost-benefit analysis should be conducted. Currently in the UK, many clinical commissioning groups fund more than one cycle of IVF treatment. If it can be shown that the additional cost of serial ultrasound scans could be recuperated from the money saved from funding additional cycles, an AFC based treatment protocol could be introduced universally.

An extended longitudinal study of serial ultrasound scans and blood tests over many consecutive menstrual cycles could help to identify the intercycle variation better. Such a study could be used to develop algorithms which could predict a woman's reproductive potential. For instance, it is conceivable where a

woman undergoes repeated ovarian reserve testing over a defined period of time, an algorithm could be created to estimate her reproductive lifespan. This could enable couples to make informed choices about starting a family. This could be particularly helpful for those with a family history of early menopause.

Appendices

Health Research Authority

Appendix 1

Prof Arri Coomarasamy
Professor of Gynaecology
University of Birmingham
Academic Unit, 3rd Floor Birmingham Women's Hospital
Foundation Trust
Mindelsohn Way
Edgbaston
B15 2TG

Email: hra.approval@nhs.net

22 November 2016

Dear Professor Coomarasamy,

Letter of HRA Approval

Study title:	**Intercycle variation in AFC, AMH and FSH in predicted poor Responders**
IRAS project ID:	**204528**
REC reference:	**16/NS/0104**
Sponsor	**University of Birmingham**

I am pleased to confirm that **HRA Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications noted in this letter.

Participation of NHS Organisations in England

The sponsor should now provide a copy of this letter to all participating NHS organisations in England.

Appendix B provides important information for sponsors and participating NHS organisations in England for arranging and confirming capacity and capability. **Please read *Appendix B* carefully**, in particular the following sections:

- *Participating NHS organisations in England* – this clarifies the types of participating organisations in the study and whether or not all organisations will be undertaking the same activities
- *Confirmation of capacity and capability* - this confirms whether or not each type of participating NHS organisation in England is expected to give formal confirmation of capacity and capability. Where formal confirmation is not expected, the section also provides details on the time limit given to participating organisations to opt out of the study, or request additional time, before their participation is assumed.
- *Allocation of responsibilities and rights are agreed and documented (4.1 of HRA assessment criteria)* - this provides detail on the form of agreement to be used in the study to confirm capacity and capability, where applicable.

Further information on funding, HR processes, and compliance with HRA criteria and standards is also provided.

It is critical that you involve both the research management function (e.g. R&D office) supporting each organisation and the local research team (where there is one) in setting up your study. Contact details and further information about working with the research management function for each organisation can be accessed from www.hra.nhs.uk/hra-approval.

Appendices

The HRA Approval letter contains the following appendices:

- A – List of documents reviewed during HRA assessment
- B – Summary of HRA assessment

After HRA Approval

The document *"After Ethical Review – guidance for sponsors and investigators",* issued with your REC favourable opinion, gives detailed guidance on reporting expectations for studies, including:

- Registration of research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics, and is updated in the light of changes in reporting expectations or procedures.

In addition to the guidance in the above, please note the following:

- HRA Approval applies for the duration of your REC favourable opinion, unless otherwise notified in writing by the HRA.
- Substantial amendments should be submitted directly to the Research Ethics Committee, as detailed in the *After Ethical Review* document. Non-substantial amendments should be submitted for review by the HRA using the form provided on the HRA website, and emailed to hra.amendments@nhs.net.
- The HRA will categorise amendments (substantial and non-substantial) and issue confirmation of continued HRA Approval. Further details can be found on the HRA website.

Scope

HRA Approval provides an approval for research involving patients or staff in NHS organisations in England.

If your study involves NHS organisations in other countries in the UK, please contact the relevant national coordinating functions for support and advice. Further information can be found at http://www.hra.nhs.uk/resources/applying-for-reviews/nhs-hsc-rd-review/.

If there are participating non-NHS organisations, local agreement should be obtained in accordance with the procedures of the local participating non-NHS organisation.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please email the HRA at hra.approval@nhs.net. Additionally, one of our staff would be happy to call and discuss your experience of HRA Approval.

HRA Training

We are pleased to welcome researchers and research management staff at our training days – see details at http://www.hra.nhs.uk/hra-training/

Your IRAS project ID is **204528**. Please quote this on all correspondence.

Yours sincerely,

Steph Blacklock
Senior Assessor

Email: hra.approval@nhs.net

Copy to: *Dr Sean Jennings, Sponsor contact*
Mrs Kelly Hard, Birmingham Women's Hospital, Lead R&D Contact
Dr Bala Karunakaran, Student

Appendix A - List of Documents

The final document set assessed and approved by HRA Approval is listed below.

Document	*Version*	*Date*
Covering letter on headed paper [Response to Provisional Opinion]		22 September 2016
GP/consultant information sheets or letters [GP letter]	1.0	01 September 2016
Interview schedules or topic guides for participants [INCA interview]	1.1	22 September 2016
IRAS Application Form	204528/1003284/37/747	01 September 2016
Letter from sponsor		31 August 2016
Other [Insurance Certificate]		28 July 2016
Other [University Peer Review]	* date received	05 September 2016
Other [Dr Chu Peer Review]		01 July 2016
Other [HRA schedule event]	1.1	22 November 206
Other [statement of activities]	1.0	22 November 2016
Participant consent form	1.1	09 August 2016
Participant information sheet (PIS)	1.2	22 September 2016
Research protocol or project proposal [INCA protocol]	1.3	20 November 2016
Summary CV for Chief Investigator (CI) [Arri Coomarasamy]		01 September 2016

Appendix B - Summary of HRA Assessment

This appendix provides assurance to you, the sponsor and the NHS in England that the study, as reviewed for HRA Approval, is compliant with relevant standards. It also provides information and clarification, where appropriate, to participating NHS organisations in England to assist in assessing and arranging capacity and capability.

For information on how the sponsor should be working with participating NHS organisations in England, please refer to the, *participating NHS organisations*, *capacity and capability* and *Allocation of responsibilities and rights are agreed and documented (4.1 of HRA assessment criteria)* sections in this appendix.

The following person is the sponsor contact for the purpose of addressing participating organisation questions relating to the study:

Dr Bala Karunakaran

HRA assessment criteria

Section	HRA Assessment Criteria	Compliant with Standards	Comments
1.1	IRAS application completed correctly	Yes	No comments
2.1	Participant information/consent documents and consent process	Yes	No comments
3.1	Protocol assessment	Yes	Protocol has been updated to V1.3 in order to comply with HRA standards.
4.1	Allocation of responsibilities and rights are agreed and documented	Yes	Statement of Activities and Schedule of Events have been provided by sponsor for use with the participating organisation.
4.2	Insurance/indemnity arrangements assessed	Yes	Where applicable, independent contractors (e.g. General Practitioners) should ensure that the professional indemnity provided by their medical defence organisation covers the activities expected of them for this research study

Section	HRA Assessment Criteria	Compliant with Standards	Comments
4.3	Financial arrangements assessed	Yes	There is no external funding acquired for the study and as per the Statement of Activities there are no funds available for the participating organisation.
5.1	Compliance with the Data Protection Act and data security issues assessed	Yes	No comments
5.2	CTIMPS – Arrangements for compliance with the Clinical Trials Regulations assessed	Not Applicable	Not Applicable
5.3	Compliance with any applicable laws or regulations	Yes	Study complies with the Human Tissue Act.
6.1	NHS Research Ethics Committee favourable opinion received for applicable studies	Yes	No comments
6.2	CTIMPS – Clinical Trials Authorisation (CTA) letter received	Not Applicable	Not Applicable
6.3	Devices – MHRA notice of no objection received	Not Applicable	Not Applicable
6.4	Other regulatory approvals and authorisations received	Not Applicable	Not Applicable

Participating NHS Organisations in England

This provides detail on the types of participating NHS organisations in the study and a statement as to whether the activities at all organisations are the same or different.
This is a single site, basic science, student study with one site type. The aim of the study is to see whether variability in the number of follicles in the ovary exists and if so whether this can be exploited to improve fertility treatment for women. Study also aims to look at how acceptable such treatment will be for women and the cost effectiveness. Participants will undergo blood tests and interview. The Chief Investigator or sponsor should share relevant study documents with participating NHS organisations in England in order to put arrangements in place to deliver the study. The documents should be sent to both the local study team, where applicable, and the office providing the research management function at the participating organisation. For NIHR CRN Portfolio studies, the Local LCRN contact should also be copied into this correspondence. For further guidance on working with

participating NHS organisations please see the HRA website.

If chief investigators, sponsors or principal investigators are asked to complete site level forms for participating NHS organisations in England which are not provided in IRAS or on the HRA website, the chief investigator, sponsor or principal investigator should notify the HRA immediately at hra.approval@nhs.net. The HRA will work with these organisations to achieve a consistent approach to information provision.

Confirmation of Capacity and Capability

This describes whether formal confirmation of capacity and capability is expected from participating NHS organisations in England.

Participating NHS organisations in England that are recruiting participants and taking blood samples **will be expected to formally confirm their capacity and capability to host this research**.

- Following issue of this letter, participating NHS organisations in England may now confirm to the sponsor their capacity and capability to host this research, when ready to do so. How capacity and capacity will be confirmed is detailed in the *Allocation of responsibilities and rights are agreed and documented (4.1 of HRA assessment criteria)* section of this appendix.
- The Assessing, Arranging, and Confirming document on the HRA website provides further information for the sponsor and NHS organisations on assessing, arranging and confirming capacity and capability.

Principal Investigator Suitability

This confirms whether the sponsor position on whether a PI, LC or neither should be in place is correct for each type of participating NHS organisation in England and the minimum expectations for education, training and experience that PIs should meet (where applicable).

Principal Investigator identified and listed in IRAS Part C.

GCP training is not a generic training expectation, in line with the HRA statement on training expectations.

HR Good Practice Resource Pack Expectations

This confirms the HR Good Practice Resource Pack expectations for the study and the pre-engagement checks that should and should not be undertaken

All staff have either full or honorary contracts with the participating organisation and therefore no further HR accesses are required.

IRAS project ID	204528

Other Information to Aid Study Set-up

This details any other information that may be helpful to sponsors and participating NHS organisations in England to aid study set-up.
• The applicant has indicated that they do not intend to apply for inclusion on the NIHR CRN Portfolio.

Appendix 2

Birmingham Women's and Children's NHS Foundation Trust

PATIENT INFORMATION SHEET

Inter Cycle Variation in Antral Follicle Count (INCA1) Study

Thank you for reading this leaflet.

We would like to invite you to take part in a study to investigate the monthly difference in fertility markers. We will be looking at the Antral Follicle Count (AFC), which looks at the structures in your ovary that can potentially release eggs. In addition we will be looking at two different hormones in your blood. These hormones are linked to fertility and are known as Anti-mullerian Hormone (AMH) and Follicle Stimulating Hormone.
Before you decide whether to take part we would like you to understand why the research is being done and what it will involve. Please take time to read this information leaflet. .One of our team will go through the information sheet with you and answer any questions that you have.

What is the purpose of the study?

The aim of the study is to improve fertility treatments by looking at monthly changes in fertility markers. The information obtained from this study will be useful to determine how we can personalise treatments to benefit women trying to get pregnant. The information gathered will help us understand your perspective better and help us to tailor treatments to meet your needs.

Why have I been invited?

Understanding how ovaries and hormones vary between monthly cycles may give us the crucial information to provide more targeted and personalised treatments to women who are trying to get pregnant. You have been invited to take part as you have been referred to our clinic with difficulties in trying to conceive and we have identified you to have one of the risk factors for infertility that we are studying.

Do I have to take part?

It is up to you to decide if you want to participate in this study. If you agree to take part we will request you to sign a written consent form. You are entitled to withdraw from the study at any time, without having to give a reason. Your choice will not affect the standard of your medical care in any way.

What will happen if I choose to take part?

You will be asked to sign a consent form. With your consent we will inform your GP that you are participating in this study. You will be given a dedicated telephone number to contact the research team on. You will need to contact us on the first day of your menstrual cycle (first day of bleeding). We will then arrange to see you in the clinic between the first and fourth day of your menstrual cycle. At this appointment we will perform an internal vaginal ultrasound scan. During this visit, we will also take approximately 10ml or 2 tablespoon of blood to check your hormone levels. The above explained procedure will be repeated for three consecutive menstrual cycles. We will share the information with you and at your request we will be happy to share these with your GP or any other healthcare provider. During these investigations if we discover something that needs urgent attention, such as a large ovarian cyst, we will pass the information on to the relevant healthcare team, only with your permission. We plan to also conduct a interview to better understand your views about the study and cater your treatment plan to your needs. If you proceed to have fertility treatment at Birmingham Women's and Children's NHS Foundation Trust, we would also review your patient notes at a later date to collect information.
Please be ensured that we will always maintain all information collected in the study will remain strictly confidential in the same way as your other medical records

What are the disadvantages of participating in the study?

Apart the inconvenience of having blood tests and internal ultrasound scans, there is no disadvantage of participating in this study. Your fertility treatment will not be delayed by you taking part in this study.

What benefits can come from participating in the study?

If you choose to participate in this study, you will receive extra scans and blood tests (normally patients would have only had one set of these). These may help you and your clinicians to better understand and plan the treatment you require .

Who has reviewed this study?

This study has been peer reviewed for scientific validity by the University of Birmingham. The North of Scotland (1) Research Ethics Committee has reviewed this study.

For further information

You can discuss the details of this study with:

Prof. Arri Coomarasamy (Consultant Obstetrician and Gynaecologist),
Dr. Bala Karunakaran (Clinical Research Fellow)

Contact:
Email:

Alternatively you can discuss participation with:
PALS (Patient advice and liaison service)

You can talk to PALS who provide confidential advice and support to patients, families and their carers, and can provide information on the NHS and health related matters.

PALS is a "confidential, friendly listening service" for people who would like to comment on any aspect of their treatment.
PALS is sited in the Front of House area of the hospital, you can leave a message in one of our suggestion boxes.

Tel: 0121 627 2747

Appendix 3

Birmingham Women's and Children's NHS Foundation Trust

In Cycle Antral Follicle Count (INCA) Study

Participant Consent Form

Please initial each box

I confirm that I have read and understand the information sheet (v1.0, 25/07/2016) for the above study. I have had the opportunity to consider the information, ask questions and these have been answered satisfactorily. ☐

I understand that my participation is voluntary and that if I take part, I am free to withdraw at any time, without giving a reason, and without my treatment being affected. ☐

I understand what is involved in the study and will be happy have extra blood tests and transvaginal ultrasound scans. ☐

I understand that my results as well as my patient records will be viewed by the research team, who are outside of my direct care team as well as my direct care team and that all clinical information will be kept confidential. ☐

I understand that my General Practitioner or any other healthcare providers will only be notified of my participation and results with my consent. ☐

I understand that the research team may wish to contact me in order to gather further information if required. ☐

I understand that relevant data collected during the study, may be looked at by individuals from the Birmingham Women's and Children's NHS Foundation Trust, the University of Birmingham, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to this data. ☐

I agree to participate in the study. ☐

…….	……	……..
Name of Patient	Date	Signature
…….	……	……..
Name of Person taking consent	Date	Signature

Trial Number …….

Appendix 3

INCA Interview Guide

Experiences of women

Informed Consent Obtained

Participant's Questions Answered

Consent to Record Obtained

Check

1. (i) How do you feel about subfertility?

(ii) How does it affect your day to day life?

(iii)Awareness of age related decline in ovarian reserve and their views about it?

(iv)Views of decisions around family planning and decision on when to conceive?

(v) Other views and observations

2. Views on Medical Trials [General]

(2a) What do you think about medical trials…

(i)for individuals?

(ii) for medical science?

(2b) What do you think about randomisation?

(i)Understandings of randomisation/ how treatment is allocated

(ii)Is randomisation acceptable to you?

(iii) Is the possibility of not getting treatment acceptable?

3. Views on Participation in a future RCT

(3a) How do you feel about the trial? Is it a necessary/worthy trial? Why? Why not?

- NB: Expand into an open-ended discussion about the trial

(i)Hopes for the trial

(ii) Concerns about the trial

(3b) What would motivate/motivated you to take part in the trial?

(i) What would you hope to get out of participating in the trial?

(3d) Trial factors?

(i)Concerns about treatment availability

(ii)Concerns about treatment choice and randomisation

4. Experiences of INCA study

(4a) How did you feel about having serial ultrasound scans and blood tests?

(4b) What were the positive experiences?

(4c) What were the negative experiences?

5. Concluding Questions

(5a) Is there anything else we didn't discuss that you would like to talk about?

(5b) Do you have questions for me?

Thank participant

- Reinforce how valuable their participation has been

Interview Guide
Clinicians

Appendix 4

-

Check	
Consent for interview	
Consent to be recorded	
Record gender, profession, and sector	

1. (i) How do you feel treating patients with low ovarian reserve?
2. (ii) How do you feel about consultations with women with low ovarian reserve?
   (iii)What do you think about the awareness of age related decline in ovarian resrve amongst women?

   (iv) Are there currently enough treatment options for women with low ovarian reserve?

   (v) Are women with low ovarian reserve have their needs catered for where you work?

2. Views on Medical Trials [General]

(2a) What do you think about medical trials…

(i)for individuals?

(ii) for medical science?

(2b) What do you think about randomisation?

(i)Understandings of randomisation/ how treatment is allocated

(ii)Is randomisation acceptable to you?

(iii) Is the possibility of not getting treatment acceptable?

3. Views on Participation in a future RCT

(3a) How do you feel about the proposed study? Is it a necessary/worthy trial? Why? Why not?

- NB: Expand into an open-ended discussion about the trial

(i)Hopes for the trial

(ii) Concerns about the trial

(3b) What would motivate/motivated you to take part in the trial?

(i)What would you hope to get out of participating in the trial?

(3d) Trial factors?

(i)Concerns about treatment availability

(ii)Concerns about treatment choice and randomisation

4..Concluding Questions

(5a) Is there anything else we didn't discuss that you would like to talk about?

(5b) Do you have questions for me?

Thank participant

- Reinforce how valuable their participation has been

Appendix 5

Analysis using binomial regression and a generalised linear model

AFC Thresholds
0 =≤5
1= 6-10
2= 11-15
3=16-20
4=21-25
5 =>/25

Unadjusted live birth by AFC categories expressed as odds ratio.
Calculated using binomial regression and generalised linear model.

	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
afc3						
0	.4015642	.045656	-8.02	0.000	.3213492	.5018024
1	.5331628	.0398011	-8.42	0.000	.4605925	.6171671
2	.725376	.0501389	-4.64	0.000	.6334714	.8306141
3	1	(base)				
4	.9595529	.0766739	-0.52	0.605	.8204516	1.122238
5	1.119551	.075853	1.67	0.096	.9803306	1.278544

Adjusted live birth by AFC categories expressed as odds ratio.
Calculated using binomial regression and generalised linear model.

	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
0	.3661979	.0556958	-6.61	0.000	.2718034	.4933746
1	.65628	.0572456	-4.83	0.000	.5531476	.7786411
2	.7991591	.0646871	-2.77	0.006	.6819203	.9365541
3	1	(base)				
4	.9006357	.0829536	-1.14	0.256	.7518804	1.078821
5	1.022202	.0806618	0.28	0.781	.8757267	1.193177
age	.9140756	.0052646	-15.60	0.000	.9038153	.9244525
BMI	1.006336	.0067396	0.94	0.346	.9932131	1.019633
ethnic	.9589892	.0257223	-1.56	0.118	.9098767	1.010753
_cons	11.67237	2.977057	9.63	0.000	7.080419	19.24241

Unadjusted pregnancy loss by AFC categories expressed as odds ratio.
Calculated using binomial regression and generalised linear model.

	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
0	2.452394	.4005442	5.49	0.000	1.780596	3.377654
1	1.655043	.1942103	4.29	0.000	1.314999	2.083019
2	1.393445	.1531507	3.02	0.003	1.123404	1.728398
3	1	(base)				
4	1.263697	.158512	1.87	0.062	.9882626	1.615896
5	1.033236	.1128486	0.30	0.765	.8341271	1.279872
_cons	.2917342	.0247165	-14.54	0.000	.2470992	.3444318

Unadjusted early pregnancy loss by AFC categories expressed as odds ratio.
Calculated using binomial regression and generalised linear model

	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
0	2.0588	.3672437	4.05	0.000	1.451373	2.920447
1	1.190305	.1607268	1.29	0.197	.9135255	1.550944
2	1.18598	.1473253	1.37	0.170	.9296921	1.512919
3	1	(base)				
4	1.187876	.1669747	1.22	0.221	.9018229	1.564663
5	.9230527	.1143758	-0.65	0.518	.7240255	1.176791
_cons	.2021116	.0191452	-16.88	0.000	.1678651	.2433448

Adjusted early pregnancy loss by AFC categories expressed as odds ratio.
Calculated using binomial regression and generalised linear model

	Odds Ratio	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
0	2.294793	.5198625	3.67	0.000	1.472016	3.57746
1	1.175154	.1829813	1.04	0.300	.8660755	1.594535
2	1.187467	.172063	1.19	0.236	.893888	1.577466
3	1	(base)				
4	1.319734	.2149253	1.70	0.088	.9591003	1.815971
5	1.069305	.1538504	0.47	0.641	.806551	1.417659
age	1.074915	.0122651	6.33	0.000	1.051143	1.099225
BMI	1.005619	.0117067	0.48	0.630	.9829344	1.028828
ethnic	.968652	.0488965	-0.63	0.528	.8774049	1.069389
_cons	.0138957	.0068258	-8.71	0.000	.0053058	.036392

Unadjusted late pregnancy loss by AFC categories expressed as odds ratio.
Calculated using binomial regression and generalised linear model.

	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
0	2.294477	.5937006	3.21	0.001	1.381763	3.810078
1	2.476123	.4680748	4.80	0.000	1.709484	3.586571
2	1.727147	.321817	2.93	0.003	1.198743	2.488471
3	1	(base)				
4	1.338204	.2905011	1.34	0.180	.8744622	2.047875
5	1.312051	.2467056	1.44	0.149	.907609	1.896717
_cons	.0612517	.0093046	-18.38	0.000	.0454793	.0824939

Adjusted late pregnancy loss by AFC categories expressed as odds ratio.
Calculated using binomial regression and generalised linear model.

	Odds Ratio	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
0	2.247793	.6651205	2.74	0.006	1.258596	4.014453
1	1.840694	.3897261	2.88	0.004	1.21551	2.787434
2	1.43828	.2981768	1.75	0.080	.9580246	2.159286
3	1	(base)				
4	1.326771	.3173754	1.18	0.237	.830197	2.120367
5	1.275684	.268463	1.16	0.247	.8445219	1.926972
age	1.108315	.0186872	6.10	0.000	1.072287	1.145553
BMI	.9729895	.0153547	-1.74	0.083	.9433556	1.003554
ethnic	1.028827	.0643815	0.45	0.650	.9100731	1.163077
_cons	.0038042	.0026584	-7.97	0.000	.0009671	.0149653

Unadjusted early pregnancy loss by AFC categories expressed as risk ratio.
Calculated using binomial regression and generalised linear model.

```
------------------------------------------------------------------------------
             | Risk Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------

           0 |   1.747683    .231867     4.21   0.000     1.347513    2.266692
           1 |   1.153401   .1275341     1.29   0.197     .9286691    1.432516
           2 |    1.15002   .1173449     1.37   0.171     .9415671    1.404623
           3 |          1  (base)
           4 |   1.151503   .1324132     1.23   0.220     .9191445      1.4426
           5 |    .935151   .0969198    -0.65   0.518     .7632424    1.145779
             |
       _cons |   .1681305   .0132486   -22.63   0.000     .1440695    .1962098
------------------------------------------------------------------------------\
```

Unadjusted late pregnancy loss by AFC categories expressed as risk ratio.
Calculated using binomial regression and generalised linear model.

```
------------------------------------------------------------------------------
             |
             | Risk Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------

           0 |   2.134968   .4971256     3.26   0.001     1.352663    3.369714
           1 |   2.281728    .395424     4.76   0.000     1.624615    3.204625
           2 |   1.657581   .2871513     2.92   0.004     1.180366    2.327731
           3 |          1  (base)
           4 |   1.312582   .2658824     1.34   0.179     .8824771    1.952314
           5 |   1.288838   .2269589     1.44   0.150     .9126526    1.820084
             |
       _cons |   .0577164   .0082615   -19.93   0.000     .0435972    .0764083
------------------------------------------------------------------------------
```

Appendix 6

Plans for dissemination

Chapter 3: Proposal accepted by Human Reproduction Update. Manuscript with wider authorship for comments and edits.

Chapter 4: Manuscript prepared for RBM Online and awaiting comments and edits from the wider authorship

Chapter 5: Manuscript being prepared by this author for Human Reproduction.

Chapter 6 and 7: Accepted for presentation at RCOG world congress in June. No decisions taken yet on intended journal for submission.

References

1. Covington SN, Burns LH, Psychology of counselling In: Sharon N. Covington,. Infertility counseling : a comprehensive handbook for clinicians [Internet]. Washington: Cambridge University Press; 2006 [cited 2018 Nov 12].

2. Wilson JF, Kopitzke EJ. Stress and infertility. Curr Womens Health Rep [Internet]. 2002 Jun [cited 2018 Nov 12];2(3):194–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12099195

3. Oddens BJ, den Tonkelaar I, Nieuwenhuyse H. Psychosocial experiences in women facing fertility problems--a comparative survey. Hum Reprod [Internet]. 1999 Jan [cited 2018 Nov 12];14(1):255–61.

4. 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 Dec 18;9(12):e1001356.

5. Boivin J, Bunting L, Collins JA, Nygren KG. International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. Hum Reprod. 2007 Jun 1;22(6):1506–12.

6. Ombelet W. Global access to infertility care in developing countries: a case of human rights, equity and social justice. Facts, views Vis ObGyn [Internet]. 2011 [cited 2018 Dec 10];3(4):257–66. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24753875

7. Symonds IM, Arulkumaran S, Symonds EM (Edwin M, Symonds EM (Edwin M.

Essential obstetrics and gynaecology. 5th ed. Ian Symonds Sabaratnam Arulkumaran, editor. London: Churchill Livingstone/Elsevier; 2013. 265–276 p.

8. Fertility problems: assessment and treatment | Guidance and guidelines | NICE [Internet]. NICE; [cited 2018 Nov 5]. Available from: https://www.nice.org.uk/guidance/cg156

9. Biggers JD. IVF and embryo transfer: historical origin and development. Reprod Biomed Online [Internet]. 2012 Aug 1 [cited 2018 Dec 4];25(2):118–27. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1472648312002921

10. Christian De Geyter. More than 8 million babies born from IVF since the world's first in 1978 [Internet]. ESHRE press release. 2018 [cited 2018 Dec 4]. Available from: https://www.eshre.eu/ESHRE2018/Media/ESHRE-2018-Press-releases/De-Geyter.aspx

11. Balen A. Assisted conception therapies. In: Balen A, editor. Infertility in practice. 4th ed. CRC press, New York; 2003.

12. National Institute for Health and Clinical Excellence. National costing report: fertility (February 2013) Costing report [Internet]. 2013 [cited 2018 Dec 4]. Available from: www.nice.org.uk

13. De Klerk C, Macklon NS, Heijnen EMEW, Eijkemans MJC, Fauser BCJM, Passchier J, et al. The psychological impact of IVF failure after two or more cycles of IVF with a mild versus standard treatment strategy. Human Reproduction, Volume 22, Issue 9, September 2007,

14. Milazzo A, Mnatzaganian G, Elshaug AG, Hemphill SA, Hiller JE,. Astute health study group on behalf of TAHS. Depression and anxiety outcomes associated with failed assisted reproductive technologies: A systematic review and meta-analysis. PLoS One [Internet]. 2016 11(11):e0165805.

15. Howard S. The hidden costs of infertility treatment.BMJ 2018;361:k2204

16. Kjaer T, Albieri V, Jensen A, Kjaer SK, Johansen C, Dalton SO. Divorce or end of cohabitation among Danish women evaluated for fertility problems. Acta Obstet Gynecol Scand. 2014 Mar 1;93(3):269–76. Available from: http://doi.wiley.com/10.1111/aogs.12317

17. Gautam A. Complications of IVF. J Obstet Gynecol India [Internet]. 2010 Aug 13;60(4):297–8.

18. McLernon DJ, Steyerberg EW, Te Velde ER, Lee AJ, Bhattacharya S. Predicting the chances of a live birth after one or more complete cycles of in vitro fertilisation: population based study of linked cycle data from 113 873 women. BMJ ,2016 Nov 16 355:i5735.

19. Shrestha D, La X, Feng HL. Comparison of different stimulation protocols used in in vitro fertilization: a review. Ann Transl Med [Internet]. 2015 Jun;3(10):137.

20. Cohen J, Trounson A, Dawson K, Jones H, Hazekamp J, Nygren K-G, et al. The early days of IVF outside the UK. Hum Reprod Update [Internet]. 2005 Oct 1 12];11(5):439–60.

21. Badawy A, Wageah A, El Gharib M, Osman EE. Strategies for pituitary down-regulation to optimize IVF/ICSI outcome in poor ovarian responders. J Reprod Infertil

[Internet]. 2012 Jul;13(3):124–30.

22. Kumar P, Sharma A. Gonadotropin-releasing hormone analogs: Understanding advantages and limitations. J Hum Reprod Sci [Internet]. 2014 Jul;7(3):170.

23. Sunkara SK, Rittenberg V, Raine-Fenning N, Bhattacharya S, Zamora J, Coomarasamy A. Association between the number of eggs and live birth in IVF treatment: an analysis of 400 135 treatment cycles. Hum Reprod [Internet]. 2011 Jul;26(7):1768–74.

24. Zhou J, Wang B, Hu Y, Sun H. Association between the number of oocytes retrieved and cumulative live birth rate in women aged 35–40 years undergoing long GnRH agonist IVF/ICSI cycles. Arch Gynecol Obstet [Internet]. 2017 Nov 6;296(5):1005–12.

25. Rienzi L, Ubaldi F, Anniballo R, Cerulo G, Greco E. Preincubation of human oocytes may improve fertilization and embryo quality after intracytoplasmic sperm injection. Hum Reprod. 1998 Apr;13(4):1014–9.

26. Ruiz A, Remohí J, Minguez Y, Guanes PP, Simón C, Pellicer A. The role of in vitro fertilization and intracytoplasmic sperm injection in couples with unexplained infertility after failed intrauterine insemination. Fertil Steril. 1997 Jul;68(1):171–3.

27. De Placido G, Wilding M, Strina I, Alviggi E, Alviggi C, Mollo A, et al. High outcome predictability after IVF using a combined score for zygote and embryo morphology and growth rate. Hum Reprod. 2002 Sep 1;17(9):2402–9.

28. Raine-Fenning N, Deb S, Jayaprakasan K, Clewes J, Hopkisson J, Campbell B. Timing of oocyte maturation and egg collection during controlled ovarian stimulation: a randomized controlled trial evaluating manual and automated measurements of follicle

diameter. Fertil Steril. 2010 Jun 1;94(1):184–8.

29. Royston P. Multiple imputation of missing values. Stata J [Internet]. 2004 [cited 2018 Nov 5];4(3):227–41. Available from: https://econpapers.repec.org/article/tsjstataj/v_3a4_3ay_3a2004_3ai_3a3_3ap_3a227-241

30. LeMaire WJ. Mechanism of mammalian ovulation. Steroids. 1989 Nov;54(5):455–69.

31. Orvieto R. Triggering final follicular maturation--hCG, GnRH-agonist or both, when and to whom? J Ovarian Res. 2015 Aug 21 8:60.

32. Humaidan P, Kol S, Papanikolaou E, Copenhagen. GnRH Agonist Triggering Workshop Group. GnRH agonist for triggering of final oocyte maturation: time for a change of practice? Hum Reprod Update. 2011 Jul 1;17(4):510–24.

33. Frattarelli JL, McWilliams GDE, Hill MJ, Miller KA, Scott RT. Low-dose aspirin use does not improve in vitro fertilization outcomes in poor responders. Fertil Steril. 2008 May 1;89(5):1113–7.

34. Ferraretti AP, La Marca A, Fauser BCJM, La Marca A ,Tarlatzis B, et al. ESHRE consensus on the definition of “poor response” to ovarian stimulation for in vitro fertilization: the Bologna criteria. Hum Reprod. 2011 Jul 1 [;26(7):1616–24.

35. Broekmans FJM, de Ziegler D, Howles CM, Gougeon A, Trew G, Olivennes F. The antral follicle count: practical recommendations for better standardization. Fertil Steril. 2010 Aug;94(3):1044–51.

36. George K, Kamath MS. Fertility and age. J Hum Reprod Sci. 2010 Sep;3(3):121–3. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21234171

37. Kim HH. Markers of ovarian reserve: is it possible to estimate an ovarian age? Fertil Steril. 2017 Dec 1;108(6):950–1.

38. Behbehani S, Buckett W, Son W, Hasson Y. Do diminished ovarian reserve markers necessarily reflect lower oocyte quality in young patients? the effect of antral follicle count. Fertil Steril. 2016 Sep 1;106(3):e124.

39. Practice Committee of the American Society for Reproductive Medicine. Testing and interpreting measures of ovarian reserve: a committee opinion. Fertil Steril 2012 Dec;98(6):1407–15.

40. Wilkosz P, Greggains GD, Tanbo TG, Fedorcsak P. Female reproductive decline is determined by remaining ovarian reserve and age. PLoS One. 2014 Oct 13;9(10):e108343.

41. Badawy A, Wageah A, El Gharib M, Osman EE. Prediction and diagnosis of poor ovarian response: the dilemma. J Reprod Infertil. 2011 Oct;12(4):241–8.

42. Zheng H, Chen S, Du H, Ling J, Wu Y, Liu H, et al. Ovarian response prediction in controlled ovarian stimulation for IVF using anti-Müllerian hormone in Chinese women. Medicine (Baltimore) [Internet]. 2017 Mar [cited 2018 Dec 28];96(13):e6495. Available from: http://insights.ovid.com/crossref?an=00005792-201703310-00043

43. Veleva Z, Jarvela I, Nuojuahuttunen S, Martikainen H and Tapanainen J. An initial low response predicts poor outcome in in vitro fertilization/intracytoplasmic sperm injection

despite improved ovarian response in consecutive cycles. Fertil Steril. 2005 May 1;83(5):1384–90.

44. Broekmans FJ, Kwee J, Hendriks DJ, Mol BW, Lambalk CB. A systematic review of tests predicting ovarian reserve and IVF outcome. Hum Reprod Update. 2006 Dec 1;12(6):685–718.

45. Raju GAR, Chavan R, Deenadayal M, Gunasheela D, Gutgutia R, Haripriya G, et al. Luteinizing hormone and follicle stimulating hormone synergy: A review of role in controlled ovarian hyper-stimulation. J Hum Reprod Sci . 2013 Oct ;6(4):227–34.

46. Sieck GC. Physiology of aging. J Appl Physiol. 2003 Oct ;95(4):1333–4.

47. Esposito MA, Coutifaris C, Barnhart KT. A moderately elevated day 3 FSH concentration has limited predictive value, especially in younger women. Hum Reprod. 2002 Jan;17(1):118–23.

48. Rajpert-De Meyts E, Jørgensen N, Græm N, Müller J, Cate RL, Skakkebæk NE. Expression of anti-müllerian hormone during normal and pathological gonadal development: association with differentiation of sertoli and granulosa cells . J Clin Endocrinol Metab . 1999 Oct;84(10):3836–44.

49. de Vet A, Laven JSE, de Jong FH, Themmen APN, Fauser BCJM. Antimüllerian hormone serum levels: a putative marker for ovarian aging. Fertil Steril. 2002 Feb;77(2):357–62.

50. Barad DH, Weghofer A, Gleicher N. Comparing anti-Müllerian hormone (AMH) and follicle-stimulating hormone (FSH) as predictors of ovarian function. Fertil Steril 2009

Apr ;91(4):1553–5.

51. Fanchin R, Taieb J, Lozano DHM, Ducot B, Frydman R, Bouyer J. High reproducibility of serum anti-Müllerian hormone measurements suggests a multi-staged follicular secretion and strengthens its role in the assessment of ovarian follicular status. Hum Reprod. 2005 Apr 1;20(4):923–7.

52. Seifer DB, MacLaughlin DT, Christian BP, Feng B, Shelden RM. Early follicular serum müllerian-inhibiting substance levels are associated with ovarian response during assisted reproductive technology cycles. Fertil Steril. 2002 Mar;77(3):468–71.

53. Cook CL, Siow Y, Taylor S, Fallat ME. Serum müllerian-inhibiting substance levels during normal menstrual cycles. Fertil Steril. 2000 Apr ;73(4):859–61.

54. Wang Y, Teng Z, Li G, Mu X, Wang Z, Feng L, et al. Cyclic AMP in oocytes controls meiotic prophase I and primordial folliculogenesis in the perinatal mouse ovary. Development. 2015 Jan 15;142(2):343–51.

55. McGee EA, Hsueh AJW. Initial and cyclic recruitment of ovarian follicles . Endocr Rev. 2000 Apr;21(2):200–14.

56. Fritz MA, Speroff L. Clinical gynecologic endocrinology and infertility. 4th ed. Wolters Kluwer Health/Lippincott Williams & Wilkins; 1989.

57. Zeleznik AJ. The physiology of follicle selection. Reprod Biol Endocrinol. 2004 Jun 16;2(1):31.

58. Young JM, McNeilly AS. Theca: the forgotten cell of the ovarian follicle. Reproduction.

2010 Oct;140(4):489–504.

59. Simoni M, Gromoll J, Nieschlag E. The follicle-stimulating hormone receptor: biochemistry, molecular biology, physiology, and pathophysiology . Endocr Rev. 1997 Dec;18(6):739–73.

60. Lai Q, Chen C, Zhang Z, Zhang S, Yu Q, Yang P, et al. The significance of antral follicle size prior to stimulation in predicting ovarian response in a multiple dose GnRH antagonist protocol. Int J Clin Exp Pathol. 2013;6(2):258–66.

61. Schipper I, de Jong FH, Fauser BC. Lack of correlation between maximum early follicular phase serum follicle stimulating hormone concentrations and menstrual cycle characteristics in women under the age of 35 years. Hum Reprod. 1998 Jun;13(6):1442–8.

62. Jayaprakasan K, Campbell B, Hopkisson J, Clewes J, Johnson I, Raine-Fenning N. Establishing the intercycle variability of three-dimensional ultrasonographic predictors of ovarian reserve. Fertil Steril. 2008 Dec 1;90(6):2126–32.

63. La Marca A, Spada E, Sighinolfi G, Argento C, Tirelli A, Giulini S, et al. Age-specific nomogram for the decline in antral follicle count throughout the reproductive period. Fertil Steril. 2011 Feb 1;95(2):684–8.

64. Nelson SM, Messow MC, Wallace AM, Fleming R, McConnachie A. Nomogram for the decline in serum antimüllerian hormone: a population study of 9,601 infertility patients. Fertil Steril. 2011 Feb;95(2):736-741.e3.

65. Hassold T, Hunt P. To err (meiotically) is human: the genesis of human aneuploidy. Nat

Rev Genet. 2001 Apr;2(4):280–91.

66. Volarcik K, Sheean L, Goldfarb J, Woods L, Abdul-Karim FW, Hunt P. The meiotic competence of in-vitro matured human oocytes is influenced by donor age: evidence that folliculogenesis is compromised in the reproductively aged ovary. Hum Reprod. 1998 Jan;13(1):154–60.

67. Greaney J, Wei Z, Homer H. Regulation of chromosome segregation in oocytes and the cellular basis for female meiotic errors. Hum Reprod Update. 2018 Mar 1;24(2):135–61.

68. Su HI, Sammel MD, Freeman EW, Lin H, DeBlasis T, Gracia CR. Body size affects measures of ovarian reserve in late reproductive age women. Menopause. 2008;15(5):857–61.

69. Freeman E, Gracia C, Sammel M, Lin H, Lim L and Strauss II. Association of anti-mullerian hormone levels with obesity in late reproductive-age women. Fertil Steril. 2007 Jan 1 [;87(1):101–6.

70. De Pergola G, Maldera S, Tartagni M, Pannacciulli N, Loverro G, Giorgino R. Inhibitory Effect of Obesity on Gonadotropin, Estradiol, and Inhibin B Levels in Fertile Women*. Obesity. 2006 Nov;14(11):1954–60.

71. Shaw ND, Histed SN, Srouji SS, Yang J, Lee H, Hall JE. Estrogen negative feedback on gonadotropin secretion: evidence for a direct pituitary effect in women. J Clin Endocrinol Metab . 2010 Apr;95(4):1955–61.

72. Schuh-Huerta SM, Johnson NA, Rosen MP, Sternfeld B, Cedars MI, Reijo Pera RA.

Genetic markers of ovarian follicle number and menopause in women of multiple ethnicities. Hum Genet. 2012 Nov;131(11):1709–24.

73. Tal R, Seifer DB. Potential mechanisms for racial and ethnic differences in antimüllerian hormone and ovarian reserve. Int J Endocrinol. 2013 Nov 21;2013:818912.

74. Gleicher N, Weghofer A, Barad DH. Ovarian reserve determinations suggest new function of FMR1 (fragile X gene) in regulating ovarian ageing. Reprod Biomed Online. 2010 Jun 1;20(6):768–75.

75. Sowers MR, McConnell D, Yosef M, Jannausch ML, Harlow SD, Randolph JF. Relating smoking, obesity, insulin resistance, and ovarian biomarker changes to the final menstrual period. Ann N Y Acad Sci. 2010 Aug;1204(1):95–103.

76. Plante BJ, Cooper GS, Baird DD, Steiner AZ. The impact of smoking on antimüllerian hormone levels in women aged 38 to 50 years. Menopause. 2010 Jan:1.

77. Nakayama M, Kamisawa E, Kawauchi H, Asada Y. The effect of lifestyle factors on anti-mullerian hormone (AMH) levels in infertile Japanese women. Fertil Steril. 2014 Sep 1;102(3):e107.

78. Holte J, Brodin T, Berglund L, Hadziosmanovic N, Olovsson M, Bergh T. Antral follicle counts are strongly associated with live-birth rates after assisted reproduction, with superior treatment outcome in women with polycystic ovaries. Fertil Steril. 2011 Sep 1;96(3):594–9.

79. Brodin T, Hadziosmanovic N, Berglund L, Olovsson M, Holte J. Antimüllerian hormone levels are strongly associated with live-birth rates after assisted reproduction. J Clin Endocrinol Metab. 2013 Mar 1;98(3):1107–14.

80. Munch EM, Sparks AE, Zimmerman MB, Van Voorhis BJ, Duran EH. High FSH dosing is associated with reduced live birth rate in fresh but not subsequent frozen embryo transfers. Hum Reprod. 2017 Jul 1 32(7):1402–9.

81. Steiner AZ, Pritchard D, Stanczyk FZ, Kesner JS, Meadows JW, Herring AH, et al. Association between biomarkers of ovarian reserve and infertility among older women of reproductive Age. JAMA. 2017 Oct 10;318(14):1367.

82. Barreto Melo MA, Garrido N, Alvarez C, Bellver J, Meseguer M, Pellicer A, et al. Antral follicle count (AFC) can be used in the prediction of ovarian response but cannot predict the oocyte/embryo quality or the in vitro fertilization outcome in an egg donation program. Fertil Steril. 2009 Jan;91(1):148–56.

83. Fang T, Su Z, Wang L, Yuan P, Li R, Ouyang N, et al. Predictive value of age-specific FSH levels for IVF-ET outcome in women with normal ovarian function. Reprod Biol Endocrinol. 2015 Jun 17;13:63.

84. Moon KY, Kim H, Lee JY, Lee JR, Jee BC, Suh CS, et al. Nomogram to predict the number of oocytes retrieved in controlled ovarian stimulation. Clin Exp Reprod Med. 2016 Jun;43(2):112–8.

85. Warburton D. The effect of maternal age on the frequency of trisomy: change in meiosis or in utero selection? Prog Clin Biol Res 1989;311:165–81.

86. Broer SL, van Disseldorp J, Broeze KA, Dolleman M, Opmeer BC, Bossuyt P, et al. Added value of ovarian reserve testing on patient characteristics in the prediction of ovarian response and ongoing pregnancy: an individual patient data approach. Hum Reprod Update. 2013 Jan 1;19(1):26–36.

87. Check JH, Summers-Chase D, Yuan W, Horwath D, Wilson C. Effect of embryo quality on pregnancy outcome following single embryo transfer in women with a diminished egg reserve. Fertil Steril. 2007 Apr 1;87(4):749–56.

88. van Disseldorp J, Lambalk CB, Kwee J, Looman CWN, Eijkemans MJC, Fauser BC, et al. Comparison of inter- and intra-cycle variability of anti-Mullerian hormone and antral follicle counts. Hum Reprod. 2010 Jan 1;25(1):221–7.

89. Scott RT, Hofmann GE, Oehninger S, Muasher SJ. Intercycle variability of day 3 follicle-stimulating hormone levels and its effect on stimulation quality in in vitro fertilization. Fertil Steril 1990 Aug;54(2):297–302.

90. Petraglia F, Sutton S, Vale W, Plotsky P. Corticotropin-releasing factor decreases plasma luteinizing hormone levels in female rats by inhibiting gonadotropin-releasing hormone release into hypophysial-portal circulation. Endocrinology. 1987 Mar;120(3):1083–8.

91. Ferin M. Stress and the Reproductive Cycle. J Clin Endocrinol Metab. 1999 Jun 1;84(6):1768–74.

92. Fertilisation H, Authority E. Fertility treatment 2014-2016 [Internet]. 2014 [cited 2018 Nov 12]. Available from: www.hfea.gov.uk

93. Arslan M, Bocca S, Mirkin S, Barroso G, Stadtmauer L, Oehninger S. Controlled ovarian hyperstimulation protocols for in vitro fertilization: two decades of experience after the birth of Elizabeth Carr. Fertil Steril. 2005 Sep;84(3):555–69.

94. Bosch E, Ezcurra D. Individualised controlled ovarian stimulation (iCOS): maximising success rates for assisted reproductive technology patients. Reprod Biol Endocrinol. 2011 Jun 21 ;9(1):82.

95. Lee T-H, Liu C-H, Huang C-C, Hsieh K-C, Lin P-M, Lee M-S. Impact of female age and male infertility on ovarian reserve markers to predict outcome of assisted reproduction technology cycles. Reprod Biol Endocrinol. 2009 Sep 17;7:100.

96. Faddy MJ, Gosden RG, Gougeon A, Richardson SJ, Nelson JF. Accelerated disappearance of ovarian follicles in mid-life: implications for forecasting menopause. Hum Reprod. 1992 Nov;7(10):1342–6.

97. Tal R, Tal O, Seifer BJ, Seifer DB. Antimüllerian hormone as predictor of implantation and clinical pregnancy after assisted conception: a systematic review and meta-analysis. Fertil Steril. 2015 Jan ;103(1):119-130.e3.

98. Deb S, Campbell BK, Clewes JS, Pincott-Allen C, Raine-Fenning NJ. Intracycle variation in number of antral follicles stratified by size and in endocrine markers of ovarian reserve in women with normal ovulatory menstrual cycles. Ultrasound Obstet Gynecol . 2013 Feb;41(2):216–22.

99. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that

evaluate healthcare interventions: explanation and elaboration. BMJ. 2009 Jul 21 ;339:b2700.

100. Tal R, Seifer DB. Ovarian reserve testing: a user's guide. 2017 [cited 2018 Nov 30]; Available from: http://dx.doi.org/10.1016/j.ajog.2017.02.027

101. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. Eur J Epidemiol. 2010 Sep 22;25(9):603–5.

102. Newcastle - Ottawa Quality Assessment Scale Cohort Studies. 2009;[Internet] [cited nov 2018] Available from : http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

103. Rustamov O, Pemberton PW, Roberts SA, Smith A, Yates AP, Patchava SD, et al. The reproducibility of serum anti-Müllerian hormone in subfertile women: within and between patient variability. Fertil Steril. 2011 Mar 1;95(3):1185–7.

104. Brown JR, Liu HC, Sewitch KF, Rosenwaks Z, Berkeley AS. Variability of day 3 follicle-stimulating hormone levels in eumenorrheic women. J Reprod Med. 1995 Sep;40(9):620–4.

105. Elter K, Si˙smanoglu A, Durmusoglu F. Intercycle variabilities of basal antral follicle count and ovarian volume in subfertile women and their relationship to reproductive aging: A prospective study. Gynecol Endocrinol. 2005 Mar 7;20(3):137–43.

106. Escobar JC, Patel SS, Carr BR, Doody KD, Doody KJ. Serum Anti-Müllerian Hormone (AMH) as a Predictor of Ovarian Response and Individual Cycle Outcome in Women

Undergoing Repetitive Egg Donation Cycles. Fertil Steril. 2010 Mar 15;93(5):S5–S5.

107. Bancsi LFJM., Broekmans FJ., Looman CW., Habbema JDF, te Velde ER. Impact of repeated antral follicle counts on the prediction of poor ovarian response in women undergoing in vitro fertilization. Fertil Steril. 2004;81(1):35–41.

108. Scheffer GJ, Broekmans FJ, Dorland M, Habbema JD, Looman CW, te Velde ER. Antral follicle counts by transvaginal ultrasonography are related to age in women with proven natural fertility. Fertil Steril. 1999 Nov ;72(5):845–51.

109. Panichkitkosolkul W. Confidence Intervals for the Coefficient of variation in a normal distribution with a known population mean. J Probab Stat. 2013 Nov 21;2013:1–11.

110. Myles PS, Cui J. Using the Bland–Altman method to measure agreement with repeated measures. Br J Anaesth. 2007 Sep 1;99(3):309–11.

111. Herberich E, Sikorski J, Hothorn T. A Robust Procedure for Comparing Multiple Means under Heteroscedasticity in Unbalanced Designs. Rapallo F, editor. PLoS One. 2010 Mar 29;5(3):e9788.

112. Fleming R, Seifer DB, Frattarelli JL, Ruman J. Assessing ovarian response: antral follicle count versus anti-Müllerian hormone. Reprod Biomed Online. 2015 Oct;31(4):486–96.

113. Müller R, Büttner P. A critical discussion of intraclass correlation coefficients. Stat Med.;13(23–24):2465–76.

114. Van Rooij IAJ, Broekmans FJM, Scheffer GJ, Looman CWN, Dik J, Habbema F, et al.

Serum antimüllerian hormone levels best reflect the reproductive decline with age in normal women with proven fertility: a longitudinal study.Fertil Steril. 2005 Apr;83(4):979-87

115. Luk J, Arici A. Does the ovarian reserve decrease from repeated ovulation stimulations? Curr Opin Obstet Gynecol. 2010 Jun;22(3):177–82.

116. Nelson SM, La Marca A. The journey from the old to the new AMH assay: how to avoid getting lost in the values. Reprod Biomed Online [Internet]. 2011 Oct [cited 2019 May 8];23(4):411–20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21872528

117. Kumar A, Kalra B, Patel A, McDavid L, Roudebush WE. Development of a second generation anti-Müllerian hormone (AMH) ELISA. J Immunol Methods. 2010 Oct 31;362(1–2):51–9.

118. Craciunas L, Roberts SA, Yates AP, Smith A, Fitzgerald C, Pemberton PW. Modification of the Beckman-Coulter second-generation enzyme-linked immunosorbent assay protocol improves the reliability of serum antimüllerian hormone measurement. Fertil Steril. 2015 Feb;103(2):554-559.e1.

119. Hendriks DJ, Mol B-WJ, Bancsi LFJMM, te Velde ER, Broekmans FJM. Antral follicle count in the prediction of poor ovarian response and pregnancy after in vitro fertilization: A meta-analysis and comparison with basal follicle-stimulating hormone level. Fertil Steril. 2005 Feb ;83(2):291–301.

120. Iliodromiti S, Kelsey TW, Wu O, Anderson RA, Nelson SM. The predictive accuracy of anti-Müllerian hormone for live birth after assisted conception: a systematic review

and meta-analysis of the literature. Hum Reprod Update. 2014 Jul 1;20(4):560–70.

121. Jurkovic D, Overton C, Bender-Atik R. Diagnosis and management of first trimester miscarriage. BMJ. 2013 Jun 19;346:f3676.

122. Simó C, Landeras J, Zuzuarregui JL, Martín JC, Remohí J, Pellicer A. Early pregnancy losses in in vitro fertilization and oocyte donation. Fertility and steril, 1999 Dec;72(6):1061-5.

123. Ezra Y, Schenker JG. Abortion rate in assisted reproduction--true increase? Early Pregnancy. 1995 Sep;1(3):171–5.

124. Lok IH, Yip AS-K, Lee DT-S, Sahota D, Chung TK-H. A 1-year longitudinal study of psychological morbidity after miscarriage. Fertil Steril. 2010 Apr;93(6):1966–75.

125. Kong G, Chung T, Lai B, Lok I. Gender comparison of psychological reaction after miscarriage-a 1-year longitudinal study. BJOG An Int J Obstet Gynaecol. 2010 Sep 1;117(10):1211–9.

126. Cheung C, Chan C, Ng E. Stress and anxiety-depression levels following first-trimester miscarriage: a comparison between women who conceived naturally and women who conceived with assisted reproduction. BJOG An Int J Obstet Gynaecol. 2013 Aug;120(9):1090–7.

127. Gonda KJ, Domar AD, Gleicher N, Marrs RP. Insights from clinical experience in treating IVF poor responders. Reprod Biomed Online. 2018 Jan 1;36(1):12–9.

128. Huyser C, Fourie FLR, Pentz J, Hurter P. The predictive value of basal follicle stimulating and growth hormone levels as determined by immunofluorometry during

assisted reproduction. J Assist Reprod Genet. 1995 Apr;12(4):244–51.

129. Tremellen K, Kolo M. Serum anti-Mullerian hormone is a useful measure of quantitative ovarian reserve but does not predict the chances of live-birth pregnancy. Aust New Zeal J Obstet Gynaecol. 2010 Dec;50(6):568–72.

130. Abdalla H, Thum MY. Repeated testing of basal FSH levels has no predictive value for IVF outcome in women with elevated basal FSH. Hum Reprod. 2006 Jan 1;21(1):171–4.

131. Bishop LA, Richter KS, Patounakis G, Andriani L, Moon K, Devine K. Diminished ovarian reserve as measured by means of baseline follicle-stimulating hormone and antral follicle count is not associated with pregnancy loss in younger in vitro fertilization patients. Fertil Steril. 2017 Dec;108(6):980–7.

132. Caroppo E, Matteo M, Schonauer LM, Vizziello G, Pasquadibisceglie A, Vitti A, et al. Basal FSH concentration as a predictor of IVF outcome in older women undergoing stimulation with GnRH antagonist. Reprod Biomed Online. 2006 Dec;13(6):815–20.

133. Chuang CC, Chen C-D, Chao K-H, Chen S-U, Ho H-N, Yang Y-S. Age is a better predictor of pregnancy potential than basal follicle-stimulating hormone levels in women undergoing in vitro fertilization. Fertil Steril. 2003 Jan;79(1):63–8.

134. Luna M, Grunfeld L, Mukherjee T, Sandler B, Copperman AB. Moderately elevated levels of basal follicle-stimulating hormone in young patients predict low ovarian response, but should not be used to disqualify patients from attempting in vitro fertilization. Fertil Steril. 2007 Apr;87(4):782-7.

135. Sabatini L, Zosmer A, Hennessy EM, Tozer A, Al-Shawaf T. Relevance of basal serum FSH to IVF outcome varies with patient age. Reprod Biomed Online [Internet]. 2008 Jul [cited 2019 Jan 31];17(1):10–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18616884

136. Thum MY, Kalu E, Abdalla H. Elevated basal FSH and embryo quality: lessons from extended culture embryos: raised FSH and blastocyst quality. J Assist Reprod Genet. 2009 Jun;26(6):313–8.

137. Toner JP, Philput CB, Jones GS, Muasher SJ. Basal follicle-stimulating hormone level is a better predictor of in vitro fertilization performance than age. Fertil Steril]. 1991 Apr;55(4):784–91.

138. Hsu A, Arny M, Knee AB, Bell C, Cook E, Novak AL, et al. Antral follicle count in clinical practice: analyzing clinical relevance. Fertil Steril. 2011 Feb 1;95(2):474–9.

139. Keane K, Cruzat VF, Wagle S, Chaudhary N, Newsholme P, Yovich J. Specific ranges of anti-Mullerian hormone and antral follicle count correlate to provide a prognostic indicator for IVF outcome. Reprod Biol. 2017 Mar;17(1):51–9.

140. Mustafa KB, Keane KN, Walz NL, Mitrovic KI, Hinchliffe PM, Yovich JL. Live birth rates are satisfactory following multiple IVF treatment cycles in poor prognosis patients. Reprod Biol. 2017 Mar;17(1):34–41.

141. Friden B, Sjoblom P, Menezes J. Using anti-Müllerian hormone to identify a good prognosis group in women of advanced reproductive age. Aust New Zeal J Obstet Gynaecol. 2011 Oct;51(5):411–5.

142. Lekamge DN, Barry M, Kolo M, Lane M, Gilchrist RB, Tremellen KP. Anti-Müllerian hormone as a predictor of IVF outcome. Reprod Biomed Online. 2007 May;14(5):602–10.

143. Pereira N, Setton R, Petrini AC, Lekovich JP, Elias RT, Spandorfer SD. Is Anti-Müllerian Hormone Associated with IVF Outcomes in Young Patients with Diminished Ovarian Reserve? Women's Heal. 2016 Mar;12(2):185–92.

144. Reijnders IF, Nelen WLDM, IntHout J, van Herwaarden AE, Braat DDM, Fleischer K. The value of Anti-Müllerian hormone in low and extremely low ovarian reserve in relation to live birth after in vitro fertilization. Eur J Obstet Gynecol Reprod Biol. 2016 May;200:45–50.

145. Tarasconi B, Tadros T, Ayoubi J-M, Belloc S, de Ziegler D, Fanchin R. Serum antimüllerian hormone levels are independently related to miscarriage rates after in vitro fertilization–embryo transfer. Fertil Steril. 2017 Sep;108(3):518–24.

146. Atasever M, Soyman Z, Demirel E, Gencdal S, Kelekci S. Diminished ovarian reserve: is it a neglected cause in the assessment of recurrent miscarriage? A cohort study. Fertil Steril. 2016 May 1;105(5):1236–40.

147. Bussen S, Sütterlin M, Steck T. Endocrine abnormalities during the follicular phase in women with recurrent spontaneous abortion. Hum Reprod. 1999 Jan 1;14(1):18–20.

148. Gürbüz B, Yalti S, Fiçicioğlu C, Ozden S, Yildirim G, Sayar C. Basal hormone levels in women with recurrent pregnancy loss. Gynecol Endocrinol. 2003 Aug ;17(4):317–21.

149. Hofmann GE, Khoury J, Thie J. Recurrent pregnancy loss and diminished ovarian

reserve. Fertil Steril. 2000 Dec 1 74(6):1192–5.

150. Pils S, Promberger R, Springer S, Joura E, Ott J. Decreased Ovarian Reserve Predicts Inexplicability of Recurrent Miscarriage? A Retrospective Analysis. Schlatt S, editor. PLoS One. 2016 Sep 14;11(9):e0161606.

151. Prakash A, Li T, Laird S, Nargund G, Ledger W. Absence of follicular phase defect in women with recurrent miscarriage. Fertil Steril. 2006 Jun;85(6):1784–90.

152. Trout SW, Seifer DB. Do women with unexplained recurrent pregnancy loss have higher day 3 serum FSH and estradiol values? Fertil Steril. 2000 Aug 1;74(2):335–7.

153. Yuan X, Lin HY, Wang Q, Li TC. Is premature ovarian ageing a cause of unexplained recurrent miscarriage? J Obstet Gynaecol (Lahore). 2012 Jul 4;32(5):464–6.

154. Eiben B, Bartels I, Bähr-Porsch S, Borgmann S, Gatz G, Gellert G, et al. Cytogenetic analysis of 750 spontaneous abortions with the direct-preparation method of chorionic villi and its implications for studying genetic causes of pregnancy wastage. Am J Hum Genet. 1990 Oct;47(4):656–63.

155. Liss J, Kunicki M, Czyzyk A, Pastuszek E, Zabielska J, Meczekalski B, et al. Clinical utility of different anti-Müllerian hormone – AMH assays for the purpose of pregnancy prediction. Gynecol Endocrinol. 2017 Oct 3;33(10):791–6.

156. Barnhart KT. Live birth is the correct outcome for clinical trials evaluating therapy for the infertile couple. Fertil Steril. 2014 May;101(5):1205–8.

157. Li HWR, Lee VCY, Lau EYL, Yeung WSB, Ho PC, Ng EHY. Role of Baseline Antral

Follicle Count and Anti-Mullerian Hormone in Prediction of Cumulative Live Birth in the First In Vitro Fertilisation Cycle: A Retrospective Cohort Analysis. PLoS One. 2013 Apr 23 ;8(4):e61095.

158. Leijdekkers JA, Eijkemans MJC, van Tilborg TC, Oudshoorn SC, McLernon DJ, Bhattacharya S, et al. Predicting the cumulative chance of live birth over multiple complete cycles of in vitro fertilization: an external validation study. Hum Reprod. 2018 Sep 1;33(9):1684–95.

159. Savitz DA, Hertz-Picciotto I, Poole C, Olshan AF. Epidemiologic measures of the course and outcome of pregnancy. Epidemiol Rev. 2002;24(2):91–101.

160. Farren J, Jalmbrant M, Ameye L, Joash K, Mitchell-Jones N, Tapp S, et al. Post-traumatic stress, anxiety and depression following miscarriage or ectopic pregnancy: a prospective cohort study. BMJ Open. 2016 Nov 2;6(11):e011864.

161. Gosalbez G, Nataloni R, … MA-… A meeting of, 2014 undefined. Is the quality of life in poor ovarian responders (POR) worse than normal responders? a prospective study using fertiqol. institutobernabeu.com [Internet]. [cited 2018 Dec 9]; Available from: https://www.institutobernabeu.com/uploads/ficheros/publicaciones/documentos-1/201504/documentos-1-instituto-bernabeu-is-the-quality-of-life-in-poor-ovarian-responders-por-worse-than-normal-responders-a-prospective-study-using-fertiqol.pdf

162. Hsu A, Arny M, Knee AB, Bell C, Cook E, Novak AL, et al. Antral follicle count in clinical practice: analyzing clinical relevance. Fertil Steril. 2011 Feb 1;95(2):474–9.

163. Haadsma ML, Groen H, Fidler V, Seinen LHM, Broekmans FJM, Heineman MJ, et al.

The predictive value of ovarian reserve tests for miscarriage in a population of subfertile ovulatory women. Hum Reprod. 2008 Dec 4;24(3):546–52.

164. Gallos ID, Khairy M, Chu J, Rajkhowa M, Tobias A, Campbell A, et al. Optimal endometrial thickness to maximize live births and minimize pregnancy losses: Analysis of 25,767 fresh embryo transfers. Reprod Biomed Online. 2018 Nov [cited 2019 Feb 4];37(5):542–8.

165. Wei D, Liu J-Y, Sun Y, Shi Y, Zhang B, Liu J-Q, et al. Frozen versus fresh single blastocyst transfer in ovulatory women: a multicentre, randomised controlled trial. Lancet. 2019 Mar 30;393(10178):1310–8.

166. Vuong LN, Pham TD, Dang VQ, Ho TM, Ho VNA, Norman RJ, et al. Live birth rates with a freeze-only strategy versus fresh embryo transfer: secondary analysis of a randomized clinical trial. Reprod Biomed Online. 2019 Mar 1;38(3):387–96.

167. Kolte AM, Bernardi LA, Christiansen OB, Quenby S, Farquharson RG, Goddijn M, et al. Terminology for pregnancy loss prior to viability: a consensus statement from the ESHRE early pregnancy special interest group. Hum Reprod. 2015 Mar 1;30(3):495–8.

168. Schmidt CO, Kohlmann T. When to use the odds ratio or the relative risk? Int J Public Health. 2008 Jun 2;53(3):165–7.

169. Dhillon RK, McLernon DJ, Smith PP, Fishel S, Dowell K, Deeks JJ, et al. Predicting the chance of live birth for women undergoing IVF: a novel pretreatment counselling tool. Hum Reprod. 2016 Jan 1;31(1):84–92.

170. El-Nemr A, Al-Shawaf T, Sabatini L, Wilson C, Lower AM, Grudzinskas JG. Effect of

smoking on ovarian reserve and ovarian stimulation in in-vitro fertilization and embryo transfer. Hum Reprod. 1998 Aug;13(8):2192–8.

171. Soares SR, Simon C, Remohi J, Pellicer A. Cigarette smoking affects uterine receptiveness. Hum Reprod. 2006 Sep 27;22(2):543–7.

172. Franasiak JM, Forman EJ, Hong KH, Werner MD, Upham KM, Treff NR, et al. The nature of aneuploidy with increasing age of the female partner: a review of 15,169 consecutive trophectoderm biopsies evaluated with comprehensive chromosomal screening. Fertil Steril. 2014 Mar 1;101(3):656-663.e1.

173. Horton NJ, Kim E, Saitz R. A cautionary note regarding count models of alcohol consumption in randomized controlled trials. BMC Med Res Methodol. 2007 Feb 15;7:9.

174. Fisch JD, Sher G. The antral follicle count (AFC) correlates with metaphase II oocytes and ART cycle outcome: an update. Fertil Steril 2002 Sep 1;78:S90.

175. Sahu B, Oztutrk O, Serhal P, Jayaprakasan K. Do ovarian reserve tests predict miscarriage in women undergoing assisted reproduction treatment? Eur J Obstet Gynecol Reprod Biol. 2010;153(2):181–4.

176. Grande M, Borobio V, Bennasar M, Stergiotou I, Mercadé I, Masoller N, et al. Role of ovarian reserve markers, antimüllerian hormone and antral follicle count, as aneuploidy markers in ongoing pregnancies and miscarriages. Fertil Steril. 2015 May;103(5):1221-1227.e2.

177. Jirge PR. Ovarian reserve tests. J Hum Reprod Sci. 2011 Sep;4(3):108–13.

178. Polyzos NP, Devroey P. A systematic review of randomized trials for the treatment of poor ovarian responders: is there any light at the end of the tunnel? Fertil Steril. 2011 Nov 1;96(5):1058-1061.e7.

179. Ulug U, Ben-Shlomo I, Turan E, Erden HF, Akman MA, Bahceci M. Conception rates following assisted reproduction in poor responder patients: a retrospective study in 300 consecutive cycles. Reprod Biomed Online. 2003 Jun;6(4):439–43.

180. Sun W, Stegmann BJ, Henne M, Catherino WH, Segars JH. A new approach to ovarian reserve testing. Fertil Steril. 2008 Dec 1;90(6):2196–202.

181. de La Rochebrochard E, Thonneau P. Paternal age and maternal age are risk factors for miscarriage; results of a multicentre European study. Hum Reprod. 2002 Jun 1;17(6):1649–56.

182. Cheng T-T. A Simplified Formula for Mean Difference. J Am Stat Assoc. 1944 Jun;39(226):240–2.

183. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016 Jun;15(2):155–63.

184. Shrout PE, Fleiss JL. Intraclass Correlations : Uses in Assessing Rater Reliability [Internet]. Vol. 86, Psychological Bulletin. 1979 [cited 2018 Nov 29]. Available from: https://www.aliquote.org/cours/2012_biomed/biblio/Shrout1979.pdf

185. Mcgraw KO. Forming Inferences About Some Intraclass Correlation Coefficients Intrinsic motivation View project. 1996 [cited 2018 Dec 3]; Available from: https://www.researchgate.net/publication/232568057

186. Altman DG, Bland JM. Measurement in Medicine: the Analysis of Method Comparison Studies † [Internet]. Vol. 32, The Statistician. 1983 [cited 2018 Nov 29]. Available from: http://people.stat.sfu.ca/~raltman/stat300/AltmanBland.pdf

187. Bush VJ, Janu MR, Bathur F, Wells A, Dasgupta A. Comparison of BD Vacutainer SST™ Plus Tubes with BD SST™ II Plus Tubes for common analytes. Clin Chim Acta [Internet]. 2001 Apr 1 [cited 2018 Nov 19];306(1–2):139–43. Available from: https://www.sciencedirect.com/science/article/pii/S0009898101003965?via%3Dihub

188. Cobas. The cobas ® 8000 analyzer series [Internet]. 2017 [cited 2018 Nov 19]. Available from: https://usdiagnostics.roche.com/en/resourcecenter/public/cd-cobas-8000-with-e-801-product-brochure-PP-US-10778.pdf

189. Coelho Neto MA, Ludwin A, Borrell A, Benacerraf B, Dewailly D, da Silva Costa F, et al. Counting ovarian antral follicles by ultrasound: a practical guide. Ultrasound Obstet Gynecol. 2018 Jan 1;51(1):10–20.

190. Dormagen JB, Gaarder M, Drolsum A. Standardized cine-loop documentation in abdominal ultrasound facilitates offline image interpretation. Acta radiol. 2015 Jan 1;56(1):3–9.

191. Coyne L, Jayaprakasan K, Raine-Fenning N. 3D Ultrasound in Gynecology and Reproductive Medicine. Women's Heal. 2008 Sep 1;4(5):501–16.

192. Loh S, Wang JX, Matthews CD. The influence of body mass index, basal FSH and age on the response to gonadotrophin stimulation in non-polycystic ovarian syndrome patients. Hum Reprod. 2002 May 1;17(5):1207–11.

193. Schliep KC, Mumford SL, Vladutiu CJ, Ahrens KA, Perkins NJ, Sjaarda LA, et al. Perceived stress, reproductive hormones, and ovulatory function: a prospective cohort study. Epidemiology. 2015 Mar;26(2):177–84.

194. Pirke KM, Schweiger U, Lemme W, Kreig JC, Berger M. The Influence of Dieting on the Menstrual Cycle of Healthy Young Women. J Clin Endocrinol Metab. 1985 Jun;60(6):1174–9.

195. Adriaens I, Cortvrindt R, Smitz J. Differential FSH exposure in preantral follicle culture has marked effects on folliculogenesis and oocyte developmental competence. Hum Reprod. 2004 Feb 1;19(2):398–408.

196. Hansen KR, Morris JL, Thyer AC, Soules MR. Reproductive aging and variability in the ovarian antral follicle count: application in the clinical setting. Fertil Steril. 2003;80(3):577–83.

197. Weghofer A, Margreiter M, Fauster Y, Schaetz T, Brandstetter A, Boehm D, et al. Age-specific FSH levels as a tool for appropriate patient counselling in assisted reproduction. Hum Reprod. 2005 Sep 1;20(9):2448–52.

198. Holton S, Rowe H, Fisher J. Women's Health and Their Childbearing Expectations and Outcomes: A Population-Based Survey from Victoria, Australia. Women's Heal Issues. 2011 Sep 1;21(5):366–73.

199. Lampic C, Svanberg AS, Karlström P, Tydén T. Fertility awareness, intentions concerning childbearing, and attitudes towards parenthood among female and male academics. Hum Reprod. 2006 Feb 1;21(2):558–64.

200. Blake D, Smith D, Bargiacchi A, France M, Gudex G. Fertility Awareness in Women Attending a Fertility Clinic. Aust New Zeal J Obstet Gynaecol. 1997 Aug 1;37(3):350–2.

201. Jirge PR. Poor ovarian reserve. J Hum Reprod Sci. 2016;9(2):63–9.

202. Findlay JK, Hutt KJ, Hickey M, Anderson RA. How Is the Number of Primordial Follicles in the Ovarian Reserve Established?1. Biol Reprod. 2015 Nov 1;93(5).

203. Maheshwari A, Fowler P, Bhattacharya S. Assessment of ovarian reserve—should we perform tests of ovarian reserve routinely? Hum Reprod. 2006 Nov 1;21(11):2729–35.

204. Botting B, Dunnell K. Trends in fertility and contraception in the last quarter of the 20th century. Popul Trends . 2000;(100):32–9.

205. Bretherick KL, Fairbrother N, Avila L, Harbord SHA, Robinson WP. Fertility and aging: do reproductive-aged Canadian women know what they need to know? Fertil Steril. 2010 May 1;93(7):2162–8.

206. Hammarberg K. Reasons for delaying childbearing A survey of women aged over 35 years seeking assisted reproductive technology. Vol. 34, Australian Family Physician. 2005.

207. Maheshwari A, Porter M, Shetty A, Bhattacharya S. Women's awareness and perceptions of delay in childbearing. Fertil Steril 2008 Oct 1;90(4):1036–42.

208. Adashi EY, Cohen J, Hamberger L, Jones HW, de Kretser DM, Lunenfeld B, et al. Public perception on infertility and its treatment: an international survey. The Bertarelli

Foundation Scientific Board. Hum Reprod. 2000 Feb;15(2):330–4.

209. O'Brien, M. Wingfield, C. Kelleher. The psychological impact of ovarian reserve testing: A qualitative analysis. Fertil Steril. 2016;106:e47.

210. Azhar E, Seifer DB, Melzer K, Ahmed A, Weedon J, Minkoff H. Knowledge of ovarian reserve and reproductive choices. J Assist Reprod Genet. 2015 Mar;32(3):409–15.

211. Bavan B, Porzig E, Baker VL. An assessment of female university students' attitudes toward screening technologies for ovarian reserve. Fertil Steril. 2011 Nov 1;96(5):1195–9.

212. Goldfarb IT, Jaffe E, James K, Lyerly AD. Pregnant women's attitudes toward Zika virus vaccine trial participation. Vaccine. 2018 Oct 29;36(45):6711–7.

213. Charlton R, De Vries C. Systematic overview of data sources for drug safety in pregnancy research [Internet]. [cited 2018 Dec 6]. Available from: http://www.encepp.eu/structure/documents/Data_sources_for_medicines_in_pregnancy_research.pdf

214. Xian Y, Wu J, O'Brien EC, Fonarow GC, Olson DM, Schwamm LH, et al. Real world effectiveness of warfarin among ischemic stroke patients with atrial fibrillation: observational analysis from Patient-Centered Research into Outcomes Stroke Patients Prefer and Effectiveness Research (PROSPER) study. BMJ. 2015 Jul 31;351:h3786.

215. Edwards V, Wyatt K, Logan S, Britten N. Consulting parents about the design of a randomized controlled trial of osteopathy for children with cerebral palsy. Heal Expect]. 2011 Dec;14(4):429–38.

216. Pollock A, Campbell P, Struthers C, Synnot A, Nunn J, Hill S, et al. Stakeholder involvement in systematic reviews: a protocol for a systematic review of methods, outcomes and effects. Res Involv Engagem. 2017;3:9.

217. Etikan I, Abubakar Musa S, Sunusi Alkassim R. Comparison of Convenience Sampling and Purposive Sampling. Am J Theor Appl Stat. 2016;5(1):1–4.

218. Patton MQ. Qualitative research & evaluation methods : integrating theory and practice. 806 p. Fourth Edition, SAGE, London

219. Suri H. Purposeful Sampling in Qualitative Research Synthesis. [cited 2018 Dec 7]; Available from: http://www.informit.com.au/qrj.html

220. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. Adm Policy Ment Health. 2015 Sep;42(5):533–44.

221. Miles MB, Huberman AM. Qualitative data analysis : an expanded sourcebook. 338 p. Second edition, 1994, SAGE, London

222. Bernard HR (Harvey R. Research methods in cultural anthropology. 1988. 520 p.Sage Publications; London

223. Boyatzis RE. Transforming qualitative information : thematic analysis and code development [Internet]. Sage Publications; 1998 [cited 2018 Dec 6]. 184 p. Available from: https://books.google.co.uk/books/about/Transforming_Qualitative_Information.html?id=_rfClWRhIKAC&redir_esc=y

224. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol. 2013 Sep 18;13:117.

225. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. BMJ . 2000 Jan 8;320(7227):114–6.

226. Heath G, Cameron E, Cummins C, Greenfield S, Pattison H, Kelly D, et al. Paediatric 'care closer to home': Stake-holder views and barriers to implementation. Health Place . 2012 Sep;18(5):1068–73.

227. Clarke V, Braun V. Successful qualitative research : a practical guide for beginners.1st edition, SAGE, London; 2013. 382 p.

228. Crabtree B, Miller W. A Template Approach to Text Analysis: Developing and Using Codebooks. Doing Qual Res Prim Care Mult Strateg [Internet]. 1992 Jan 1 [cited 2018 Dec 6]; Available from: https://scholarlyworks.lvhn.org/family-medicine/44

229. Strauss AL, Corbin JM. Basics of qualitative research : techniques and procedures for developing grounded theory [Internet]. Sage Publications; 1998 [cited 2018 Dec 6]. 312 p. Available from: https://books.google.co.uk/books/about/Basics_of_Qualitative_Research.html?id=wTwYUnHYsmMC&redir_esc=y

230. Campbell JL, Quincy C, Osserman J, Pedersen OK. Coding In-depth Semistructured Interviews: Problems of Unitization and Intercoder Reliability and Agreement. Sociol Methods Res . [cited 2018 Dec 7];42(3):294–320.

231. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Heal Care. 2007 Sep 16;19(6):349–57.

232. Anderson J, Adey P, Bevan P. Positioning place: polylogic approaches to research methodology. Qual Res. 2010 Oct 25;10(5):589–604.

233. Chew-Graham CA, May CR, Perry MS. Qualitative research and the problem of judgement: lessons from interviewing fellow professionals. Fam Pract. 2002 Jun ;19(3):285–9.

234. Mills M, Rindfuss RR, McDonald P, te Velde E, ESHRE Reproduction and Society Task Force on behalf of the ER and ST. Why do people postpone parenthood? Reasons and social policy incentives. Hum Reprod Update. 2011;17(6):848–60.

235. Newington L, Metcalfe A. Factors influencing recruitment to research: qualitative study of the experiences and perceptions of research teams. BMC Med Res Methodol. 2014 Jan 23;14:10.

236. Jenkins V, Farewell D, Batt L, Maughan T, Branston L, Langridge C, et al. The attitudes of 1066 patients with cancer towards participation in randomised clinical trials. Br J Cancer. 2010 Dec 30;103(12):1801–7.

237. Boivin J, Domar AD, Shapiro DB, Wischmann TH, Fauser BCJM, Verhaak C. Tackling burden in ART: an integrated approach for medical staff. *Human Reproduction*, Volume 27, Issue 4, April 2012, Pages 941–950

238. Harris RD, Bond MJ. Stress in IVF workers. Clin Reprod Fertil.;5(1–2):27–35.

239. Rasool S, Shah D. Fertility with early reduction of ovarian reserve: the last straw that breaks the Camel's back. Fertil Res Pract. 2017 Dec 11;3(1):15.

240. Broer SL, Broekmans FJM, Laven JSE, Fauser BCJM. Anti-Müllerian hormone: ovarian reserve testing and its potential clinical implications. Hum Reprod Update. 2014 Sep 1;20(5):688–701.

241. La Marca A, Ferraretti AP, Palermo R, Ubaldi FM. The use of ovarian reserve markers in IVF clinical practice: a national consensus. Gynecol Endocrinol. 2016 Jan 2 [;32(1):1–5.

242. Boaz A, Hanney S, Borst R, O'Shea A, Kok M. How to engage stakeholders in research: design principles to support improvement. Heal Res Policy Syst. 2018 Dec 11;16(1):60.

243. Redman S, Turner T, Davies H, Williamson A, Haynes A, Brennan S, et al. The SPIRIT Action Framework: A structured approach to selecting and testing strategies to increase the use of research in policy. Soc Sci Med. 2015 Jul;136–137:147–55.

244. Kok MO, Gyapong JO, Wolffers I, Ofori-Adjei D, Ruitenberg J. Which health research gets used and why? An empirical analysis of 30 cases. Heal Res Policy Syst. 2016 Dec 17;14(1):36.

245. Pandi-Perumal SR, Akhter S, Zizi F, Jean-Louis G, Ramasubramanian C, Edward Freeman R, et al. Project Stakeholder Management in the Clinical Research Environment: How to Do it Right. Front Psychiatry. 2015 May 18;6.

246. Esmail L, Moore E, Rein A. Evaluating patient and stakeholder engagement in research: moving from theory to practice. J Comp Eff Res. 2015 Mar;4(2):133–45.

247. Cooke J, Ariss S, Smith C, Read J. On-going collaborative priority-setting for research activity: a method of capacity building to reduce the research-practice translational gap. Heal Res Policy Syst. 2015 Dec 7;13(1):25.

248. Sharma G. Pros and cons of different sampling techniques. Int J Appl Res. 2017;3(7):749–52.

249. Boivin J, Bunting L, Koert E, ieng U C, Verhaak C. Perceived challenges of working in a fertility clinic: a qualitative analysis of work stressors and difficulties working with patients. Hum Reprod. 2017 Feb 1;32(2):403–8.

250. Peddie VL, van Teijlingen E, Bhattacharya S. A qualitative study of women's decision-making at the end of IVF treatment. Hum Reprod. 2005 Jul 1;20(7):1944–51.

251. Abu-Rabia A. Infertility and surrogacy in islamic sociert: socio-cultural, psychologicial, ethnica and religious dilemmas. Open Psychol J. 2013;6:54–60.

252. Inhorn MC. Making Muslim babies: IVF and gamete donation in Sunni versus Shi'a Islam. Cult Med Psychiatry. 2006 Dec;30(4):427–50.

253. Thoma A, Farrokhyar F, McKnight L, Bhandari M. Practical tips for surgical research: how to optimize patient recruitment. Can J Surg. 2010 Jun;53(3):205–10.

254. Mapstone J, Elbourne D, Roberts IG. Strategies to improve recruitment to research studies. In: Mapstone J, editor. Cochrane Database of Systematic Reviews [Internet]. Chichester, UK: John Wiley & Sons, Ltd; 2007 [cited 2019 Jan 20]. Available from: http://doi.wiley.com/10.1002/14651858.MR000013.pub3

255. Tishler CL, Bartholomae S. The recruitment of normal healthy volunteers: a review of the literature on the use of financial incentives. J Clin Pharmacol. 2002 Apr;42(4):365–75.

256. Culley L, Hudson N, Van Rooij F. Marginalized reproduction : ethnicity, infertility and reproductive technologies. Earthscan; 2009. 207 p.

257. Largent EA, Grady C, Miller FG, Wertheimer A. Money, coercion, and undue inducement: attitudes about payments to research participants. IRB. 2012;34(1):1–8.

258. Klitzman R. How IRBs view and make decisions about coercion and undue influence. J Med Ethics 2013 Apr;39(4):224–9.