Fertility in female cancer survivors: pathophysiology, preservation and the role of ovarian reserve testing

Kerryn Lutchman Singh1, Melanie Davies and Ratna Chatterjee

Department of Obstetrics and Gynaecology, Reproductive Medicine Unit, Elizabeth Garrett Anderson and Obstetric Hospital, University College London Hospital, Huntley Street, London WC1E 6DH, UK

1To whom correspondence should be addressed. E-mail: k.lutchman-singh@ucl.ac.uk

The improved long-term survival of adolescents and young women treated for cancer has resulted in an increased focus on the effects of chemotherapy on ovarian function and its preservation. These women may seek advice and treatment regarding their reproductive status, including ways of preserving their fertility and preventing a premature menopause—factors that can have a profound impact on their quality of life. This article comprehensively reviews ovarian reserve testing (ORT) in general. Special emphasis is placed on patients with cancer, including the pathophysiology of gonadal damage following chemotherapy, fertility preservation and the potential role of ORT. Baseline parameters of ovarian reserve [FSH LH, estradiol, inhibin B and anti-Müllerian hormone (AMH)] have not yet performed sufficiently well in predicting poor outcome in assisted reproduction, but biochemical markers of ovarian reserve appear to be better than chronological age. Inhibin B and AMH show potential for future use. Dynamic testing appears to show much promise, especially stimulated levels of inhibin B and estradiol. The most promising tests of ovarian reserve are the biophysical markers, where total antral follicle count was found to be most discriminatory followed by ovarian volume. Combination of biochemical, biophysical and clinical markers of ovarian reserve may also improve predictive capacity. However, there is a lack of data pertinent to ORT in cancer. As yet there is no single clinically useful test to predict ovarian reserve accurately. Patients with cancer represent a distinct cohort who have particular concerns about their future fertility and the possibility of a premature menopause, they can benefit greatly from knowledge of their functional ovarian reserve. Large, prospective, randomized, adequately controlled studies specific to different geographical areas are required in a control population of comparable reproductive age to determine the potential role of ORT in clinical practice.

Key words: cancer/chemotherapy/fertility/ovarian reserve tests/premature menopause

Introduction

A review of cancer trends in England and Wales from 1950 to 1999 reveals an overall increase in relative survival for adult cancers (Office for National Statistics, 2004). This includes breast cancer, the most common cancer affecting women in developed countries (Figure 1). In these patients, the decline in age-adjusted breast cancer mortality rates is almost certainly due to improvements in the earlier detection of smaller, palpable tumours and in adjuvant chemotherapy (Mettlin, 1999; Jatoi and Miller, 2003). It has been estimated that 25% of women diagnosed with breast cancer are pre-menopausal, with 7% being diagnosed before the age of 40 years (Hankey et al., 1994). Another estimate is that 10–20% of newly diagnosed breast cancers occur in women of childbearing age (Parker et al., 1997).

The remarkable success in improving childhood cancer survival is exemplified by the 5 year survival rate for all leukaemias approaching 80% (Office for National Statistics, 2004). As such, it has been estimated that at the start of the 21st century, one in 1000 young adults in their third decade is a survivor of childhood cancer (Bath et al., 2002). Overall, much of the improvement in 5 year survival rates in adult and childhood cancer is attributed to improved treatment, of which chemotherapy plays a major role (McVie, 1999).

This success, however, is tempered by the fact that chemotherapy can result in ovarian toxicity. For these patients, who originally had expectations of a normal reproductive lifespan, the realization that they might suffer reduced fertility and a premature menopause can have a profound impact on their self-esteem and quality of life (Meirow and Nugent, 2001). Furthermore, women are delaying childbearing for domestic, professional and educational reasons (Friedlander and Thewes, 2003). In England and Wales the average age of mothers at childbirth has increased by 3 years since 1971, rising from 26.2 years to 29.1 years in 2000 (Office for National Statistics, 2004).
K.Lutchman Singh, M.Davies and R.Chatterjee

Ovarian reserve testing has become established in the reproductive medicine setting, where it is utilized as a counselling tool and a guide to treatment respectively for subfertility. Ovarian ageing is a concept intimately related to ovarian reserve, as it could potentially predict age at menopause, a measure which is known to be highly variable (Treloar, 1981). Biochemical markers also show considerable variability, with FSH levels not increasing permanently until 1–5 years after the onset of the peri-menopause (Burger et al., 1999). Inhibins and activins have also been explored (Santoro et al., 1999), and it has been suggested that there is a subtle decrease in inhibin B prior to the rise of FSH (Soules et al., 1998). Ultrasound determination of ovarian volume and antral follicle count (AFC) have also been utilized with some success (Flaws et al., 2001).

Several hypotheses have been forwarded in an attempt to describe the process of ovarian ageing (Lobo, 2003; Nikolau and Templeton, 2003) but the fact remains that reproductive ageing itself is highly variable (te Velde and Pearson, 2002). The basic doctrine which has prevailed to explain the concept of ovarian reserve and reproductive ageing assumes that the age-related loss of fertility in females is dictated by a continual process of follicle depletion (which is fixed from fetal life), leading to a reduction in both oocyte quantity and quality. Even this central doctrine has recently been challenged by the reported finding of the existence of proliferative germ cells that sustain oocyte and follicle production using a mouse model (Johnson et al., 2004b).

At present therefore, it is impossible to predict the functional life span of the chemotherapeutically damaged ovary and the reproductive potential of patients with cancer. There is a need for adequate assessment of functional ovarian reserve in pre-menopausal patients with cancer, as it may provide useful information regarding their fertility status and the prediction of premature menopause. The potential also exists for these tests to act as a guide to treatment with respect to fertility preservation, where an individualized approach would be ideal.

This article comprehensively reviews ovarian reserve testing (ORT) in general. Special emphasis is placed on patients with cancer, including the pathophysiology of gonadal damage following chemotherapy, fertility preservation and the potential role of ORT.

**Pathophysiology of ovarian damage following chemotherapy**

The gonadotoxic effect of chemotherapeutic agents is well documented, although the prevailing mechanisms are not fully understood. In general, alkylating agents such as cyclophosphamide (C) which are non-cell cycle specific, are more cytotoxic to the ovaries than cell cycle-specific agents such as methotrexate (M) and fluorouracil (F), whose major effect is on ovarian follicle growth and maturation (Hensley and Reichman, 1998). Ovarian biopsies in patients undergoing cyclophosphamide-based treatment reveal complete absence of ova or small numbers of inactive ova with fibrosis and no evidence of follicular maturation (Warne et al., 1973; Koyama et al., 1977). Animal studies have shown that exposure to cyclophosphamide causes follicular destruction in exponential proportion to increasing doses (Meirow et al., 1999). These events may affect sex steroid production, leading to disturbance of the hypothalamic–pituitary–ovarian (HPO) axis (Dnistrian et al., 1985; Dowsett and Richner, 1991; Mehta et al., 1992).

Ovarian damage following chemotherapy can present with a variety of symptoms that reflect varying degrees of damage, culminating in premature ovarian failure (POF). POF is associated with a wide array of effects including vasomotor symptoms (hot flushes and night sweats), genitourinary symptoms (vaginitis, dyspareunia, dysuria) and osteoporosis, which can lead to skeletal fractures (Ganz and Greendale, 2001). Ovarian failure after adjuvant chemotherapy has been shown to be associated with rapid bone loss in women with early stage breast cancer (Shapiro et al., 2001). The management of these symptoms is particularly difficult (especially in patients with breast cancer) owing to concerns about HRT (Rostom, 2001; Beral, 2003).

Chemotherapy-related amenorrhea (CRA) is a term used to describe the occurrence of amenorrhea following chemotherapy, the rate of which varies according to the diagnostic criteria used and length of follow-up (Bines et al., 1996). Its onset is mediated through ovarian failure, and is based on the observation that the hormone profile observed in pre-menopausal women treated with adjuvant chemotherapy for breast cancer who develop CRA is consistent with primary ovarian failure (Rose and Davis, 1977; Dnistrian et al., 1983; Dowsett and Richner, 1991). The risk of POF with polyagent adjuvant chemotherapy has been reported to range from 53 to 89% (Del Mastro et al., 1997). Generally, the risk of POF is related to the patient’s age, treatment protocol and type of malignancy (Meirow and Nugent, 2001).

Age and time to CRA are inversely related, as well as cumulative dose required to produce CRA. The average incidence of CRA for CMF-based regimens is 40% for women aged <40 years and 76% for those >40 years (Bines et al., 1996). A cohort of pre-menopausal women with breast cancer, receiving either adjuvant CMF, CEF (E = epirubicin), tamoxifen, or no treatment was followed up prospectively for 1 year (Goodwin et al., 1999). Age and the use of systemic chemotherapy were found to be independent predictors of premature menopause. The use of CMF or CEF, whether in combination or not with
tamoxifen, increased the risk of menopause in women aged 40 years from <5 to >40% (Goodwin et al., 1999). This potentiating effect of age on chemotherapy-induced gonadal damage has been reported elsewhere (Sanders et al., 1988; Moore, 2000).

Restoration of menstruation after CRA is possible. Again this is influenced by age and duration of follow up, and has been estimated at 39–55% in younger women (<40 years) and 0–11% in older patients (>40 years) (Bines et al., 1996). However, women who maintain normal menses throughout chemotherapy remain at risk for developing POF. This is evident from the high rates of POF seen in adolescents receiving alkylating agents for cancer (Byrne et al., 1992).

Chatterjee and Kottaridis (2002) attempted to define this sequence of events by using a haematological model termed the gonadal insufficiency–premature gonadal failure syndrome (GI–PGF). As the name implies, the syndrome describes a heterogeneous picture whereby the gonadotoxic effect of these agents is not ‘all or nothing’ but can be both acute and cumulative, with the ovaries possessing a limited capacity for recovery. These authors made an attempt to grade the extent of damage using clinical, biochemical and biophysical parameters; however, it was noted that the clinical picture was variable and as such did not always comply with these parameters. The potential for ovarian recovery, as well as the response to therapy, was unpredictable (Chatterjee and Kottaridis, 2002).

While this model may not be directly applicable to all cancers, it remains an attractive one, which highlights the importance of developing tests that can accurately estimate ovarian damage, and correlate it with clinical features.

Effect of chemotherapy on fertility

Young cancer survivors are concerned about their reproductive capability (both before and after treatment), which can have a profound impact on their self-esteem (Ganz, 2000). Those patients who are not immediately rendered infertile are still likely to suffer problems with infertility and, ultimately, a premature menopause (Meirov, 2000). Approximately 50% of patients aged <35 years resume normal menses after completion of cytotoxic chemotherapy (Forbes, 1992), and are therefore potentially capable of becoming pregnant. However, although a regular menstrual cycle may serve as a convenient marker for ovarian function, a normal menstrual cycle is not synonymous with fertility. Similarly, irregular menses or amenorrhoea does not always imply infertility. This was shown in a retrospective series of women with breast cancer who received chemotherapy with FAC (5-fluoracil, anthracycline, cyclophosphamide) who were ≤35 years at the time of treatment. In this series, 33 pregnancies occurred in 25 patients (21%), of whom only 64% continued to menstruate regularly during and after chemotherapy (Sutton et al., 1990).

Because of the possibility of spontaneous recovery, it has been suggested that ovarian function be reassessed periodically in patients with chemotherapy-induced gonadal damage. Conversely, if fertility is not desired, contraception should be used, and the combined oral contraceptive pill (COCP) can be used for HRT (Nasir et al., 1997).

A large retrospective survey of pregnancy outcomes after peripheral blood or bone marrow transplantation revealed that only 0.6% of patients conceived after autologous or allogenic stem cell transplantation (SCT), but the pregnancies were likely to have a successful outcome (Salooja et al., 2001).

Overall, the data available relating to pregnancy rates and outcomes following cancer is limited and relates only to distinct cohorts, which may not be representative of the entire population. Furthermore, follow-up is almost always limited in duration.

Options for preserving fertility

Advances in reproductive technology have made fertility preservation techniques a real possibility for these patients. Decision-making in this area is particularly difficult because of the experimental nature of many techniques. Patients may benefit from knowledge of their functional ovarian reserve prior to embarking on these decisions.

Ovarian protection

This has been attempted by the administration of prior and concomitant treatment with GnRH analogues (Blumenfeld et al., 1999). Zoladex® has been compared with CMF where it was found that goserelin offers an effective, well-tolerated alternative to CMF in pre-menopausal patients with estrogen receptor (ER)-positive and node-positive early breast cancer (Jonat et al., 2002). Concerns exist regarding the side-effects of goserelin, but are considered to be reversible (Nystedt et al., 2003). Despite encouraging results in animal models, few clinical studies have evaluated the effect of GnRH analogue co-treatment in preventing chemotherapy induced POF in cancer patients (Blumenfeld, 2003). The questionable presence of GnRH receptors in human gonadal tissue also challenges the plausible mechanism of action (Clayton and Huhtaniemi, 1982). A randomized prospective multicentre study is currently underway in the UK attempting to address this issue (Ovarian Protection Trial In Oestrogen Non-responsive Premenopausal Breast Cancer Patients Receiving Adjuvant or Neo-adjuvant Chemotherapy: ‘OPTION’; Scottish Cancer Network).

Embryo cryopreservation

The only established method available in clinical practice to preserve fertility in women being treated for cancer is the cryopreservation of embryos prior to chemotherapy (Revel and Schenker, 2004). There are significant drawbacks to its use, namely the need for a stimulated cycle and IVF. The procedure itself may be theoretically harmful to patients with hormone-sensitive tumours such as breast cancer. This risk has been circumvented in the past by employing natural cycle IVF (Brown et al., 1996). More recently, IVF and embryo cryopreservation after ovarian stimulation with tamoxifen has been employed with some success whilst possibly providing a safe alternative to ovarian stimulation in these patients (Oktay et al., 2003). Further considerations include the fact that a harvest of viable embryos cannot be guaranteed, and any complications or delay in the IVF cycle may delay the commencement of chemotherapy (Thomson et al., 2002). The fate of the embryos should be considered at the outset, given the fact that potential offspring may lose their mothers to the disease (Posada et al., 2001).
K.Lutchman Singh, M.Davies and R.Chatterjee

Oocyte cryopreservation
This is an attractive option for women without a partner. Drawbacks are analogous to embryo cryopreservation as a stimulated cycle is required. Since the first report of a birth from a frozen oocyte (Chen, 1986), the results of this procedure worldwide have been variable with a reported success rate of <2%, despite the improved success rate when combined with ICSI (Porcu et al., 1997; Fabbri et al., 2001). Despite this, the technique is still considered to have an important place in reproductive medicine, especially in patients with cancer (Van der Elst, 2003).

Ovarian tissue cryopreservation
This has been a topic of renewed and intense interest as a viable alternative to restore both fertility and ovarian function to women with cancer (Falcone et al., 2004). Earlier studies provided a basis for the technique by showing that pregnancies and live births could be achieved in animals by reimplanting cryopreserved ovarian tissue. Although not widely practised, pregnancy rates after IVM (using immature antral follicles) have been >30% in some centres, with an estimated 300 healthy infants born worldwide (Chian et al., 2004). Primordial follicle culture however, has proven to be particularly more challenging and is the subject of intense research interest (Gosden et al., 2002).

Although the advantages of this procedure in patients with cancer is readily apparent, there are immense ethical, moral and sometimes legal dilemmas associated with these procedures and as such should not be ignored (Robertson, 2000). With this in mind, continuing research in the field has created optimism for clinical use in the future (Chian et al., 2004). Until then, it should not be considered a viable alternative for patients who want fertility preservation if more established alternatives are an option (Revel and Schenker, 2004).

Summary of fertility preservation techniques
Fertility preservation techniques have evolved to the extent that there are now several options available to the patient with cancer. It is important to remember that despite several encouraging advances made in this area, the best option for patients who are suitable remains embryo cryopreservation. With the continual development of these techniques, however, the ideal scenario may develop where it might be possible to offer an individualized approach to management (Sonmezer and Oktay, 2004).

What is ovarian reserve?
Fertility potential in the female patient is related to the total number and quality of the primordial follicles remaining in the ovaries and is referred to as ovarian reserve. The function of these follicles is 2-fold: gametogenesis, which governs fertility potential, and steroidogenesis, which governs the onset of a premature menopause.

Development of ovarian reserve tests (ORT)
ORT have become established in the fertility clinic setting where the association of poor ovarian response due to diminished reserve, leads to cycle cancellations and reduced success rates in IVF (Pellicer et al., 1987). Although natural fecundity and pregnancy rates decrease with increasing age (te Velde and Pearson, 2002), the need for ORT was clearly realized when it was established that chronological age and menstrual characteristics were unreliable in predicting reproductive age (Scott et al., 1995). Several models have now been described which accurately chart the progressive decline of primordial follicles in the ovary with a woman's age (Richardson et al., 1987; Faddy et al., 1992; Gougeon et al., 1994; Faddy and Gosden, 1996).

The most important aspect of diminished ovarian reserve and the associated decline in reproductive potential is that its onset is highly variable (Scott and Hofmann, 1995). Furthermore, the presence of regular menses does not establish the presence of adequate ovarian reserve, as the ovary has the ability to maintain a high frequency and number of ovulations despite continuously declining follicle number (Gosden, 1987). Hence the continual decline in follicle number and quality can best be described as a dynamic process, the mechanisms of which are not yet fully understood (te Velde, 1993).

Overall, most of the data pertaining to ORT are limited by a lack of prospective, controlled data outside a reproductive...
medicine setting. Therefore, until appropriate validation occurs, there is as yet no clinically useful predictive test to accurately assess ovarian reserve.

**Ovarian reserve tests**

In general, ORT are either biochemical or biophysical (Table I). Depending on the basis of the test being used, this may be further classified as being direct or indirect, or alternatively as being static or dynamic in nature.

**Biochemical markers**

**FSH**

Basal serum FSH (bFSH) is considered an indirect estimate of ovarian reserve and depends on the presence of an intact HPO axis. It is one of the longest established parameters for estimating ovarian reserve.

The first reports on the usefulness of bFSH measurements showed FSH to be predictive of ovarian response and pregnancy outcome in IVF cycles (Muasher et al., 1988; Scott et al., 1989), and to be more predictive than the patient’s age (Toner et al., 1991). Subsequent studies have supported this, including a large retrospective analysis of IVF patients where elevated FSH was associated with high pregnancy loss and poor live birth rates, regardless of age (Levi et al., 2001).

However, there are now several studies that challenge the predictive capabilities of bFSH, especially when a patient’s age is considered. For example, in patients undergoing their first IVF cycle, bFSH concentration was a better predictor of cancellation rate than age, yet age was a stronger predictor of pregnancy rate (Sharif et al., 1998; Creus et al., 2000). Some studies suggest that young age exerts a protective effect over the deleterious effects of reduced ovarian reserve (Hanoch et al., 1998), while others reveal no difference in outcome between younger and older women with elevated bFSH (El Toukhy et al., 2002). A retrospective cohort study of IVF patients revealed that increasing age, but not bFSH, was associated significantly with reduced implantation and pregnancy rates (Chuang et al., 2003). A recent prospective study, regularly cycling women aged >40 and <40 years, with elevated bFSH levels undergoing their first IVF cycle, were compared for ovarian response, ongoing pregnancy rates and implantation rates. Although younger women with elevated bFSH levels had higher cycle cancellation rates, implantation and ongoing pregnancy was superior to their older counterparts (van Rooij et al., 2003). This has prompted the distinction that when a comparison is made, age better reflects oocyte quality whereas FSH reflects oocyte quantity (Toner, 2003). This study has been criticized, however, for a lack of adequate controlling and statistical power (McDonough, 2003).

Threshold levels for bFSH are also an issue. In one study, which is unique for its duration of follow-up, ongoing pregnancy was seen in as much as 28% of regularly cycling women with FSH levels of 15–20 IU/l. Only when the FSH level was >20 IU/l was there a clear fall in ongoing pregnancy rate regardless of age (van Rooij et al., 2004). A high threshold level of FSH to achieve acceptable prediction for treatment failure has been reported elsewhere (Bancsi et al., 2000). It seems therefore that younger women with moderate elevations of FSH should not be ignored (Toner, 2004). Younger women (age <41 years) with elevated bFSH levels who have a poor first response to IVF, however, may represent a specific group which should be counselled against further attempts, as was shown in a recent retrospective study (Klinkert et al., 2004). Which threshold level of bFSH to use is the question, as some investigators have shown that moderately elevated levels of FSH (defined as 10–11.4 IU/l) are difficult to interpret as these may be confounded by a poor response to gonadotrophin stimulation (Esposito et al., 2002).

As most of the data regarding the performance of bFSH in predicting ovarian reserve comes from studies in an IVF population, it is significant that a recent meta-analysis concluded that the performance of bFSH for predicting poor ovarian response in this group was moderate, while prediction of non-pregnancy was poor (Bancsi et al., 2003). Information pertaining to the general subfertile population is lacking, but a nested case–control study revealed no statistically significant difference in cumulative pregnancy rates between patients with elevated bFSH levels and controls (Van Montfrans et al., 2000). Others have shown elevated bFSH to be of limited value in patients with regular cycles, in that it should not lead to exclusion of treatment (van Rooij et al., 2004).

Questions remain regarding the reproducibility of bFSH measurements (Sharara et al., 1998). One of its main limitations is the significant intercycle variability of FSH, hence limiting the usefulness of a normal value (Scott et al., 1990). A recent randomized prospective study revealed that women with limited ovarian reserve exhibited strong intercycle variability of bFSH and FSH response to clomiphene citrate (Kwee et al., 2004). A study performed on normally ovulating women aged <35 years revealed a large inter-individual variation in bFSH (Schipper et al., 1998). Intra-cycle variability seems to be less of an issue, with some flexibility being apparent in sampling from day 2 to day 5 of a regular menstrual cycle (Hansen et al., 1996; Klein et al., 1996a).

Another concern is the establishment of normative data within individual laboratories due to the likelihood of interassay variability (Hershlag et al., 1992). The establishment of threshold values within individual assay systems is also important to prevent errors in interpretation of results, especially when clinical outcomes are being correlated. This is more likely to happen in institutions where the clinical volume of patients is insufficient to allow these threshold values to be established (Scott, 2004).


**K.Lutchman Singh, M.Davies and R.Chatterjee**

### Table II. Basal FSH as a predictor of fertility

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Study type</td>
<td>ART</td>
<td>ART</td>
<td>General subfertile</td>
<td>General subfertile</td>
<td>ART</td>
<td>General subfertile</td>
</tr>
<tr>
<td>Patients ± cycles</td>
<td>Retrospective</td>
<td>Retrospective</td>
<td>Retrospective</td>
<td>Nested case--control</td>
<td>Prospective</td>
<td>Retrospective</td>
</tr>
<tr>
<td>Threshold (IU/l)</td>
<td>441/758</td>
<td>435</td>
<td>9802</td>
<td>100</td>
<td>86</td>
<td>122</td>
</tr>
<tr>
<td>Outcome measure</td>
<td>Ongoing PR</td>
<td>Ongoing PR</td>
<td>Pregnancy loss</td>
<td>PR Delivery Time to</td>
<td>Ongoing PR</td>
<td>Ongoing PR Time to</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rate LBR</td>
<td>pregnancy</td>
<td></td>
<td>ongoing pregnancy</td>
</tr>
<tr>
<td>Conclusion</td>
<td>bFSH predictive of pregnancy outcome</td>
<td>bFSH of limited value in predicting ongoing PR</td>
<td>bFSH predictive of pregnancy outcome</td>
<td>bFSH not predictive of pregnancy outcome</td>
<td>IVF outcome depending on age</td>
<td>bFSH of limited value in predicting outcome</td>
</tr>
</tbody>
</table>

ART = assisted reproduction treatment; PR = pregnancy rate; LBR = live birth rate; PT = patients; bFSH = basal FSH.

Other reasons for an elevated bFSH should be considered (Lambalk, 2003). Laboratory errors can occur in the presence of heterophilic antibodies, which can interfere with the FSH immunoassay (Cahill and Thomas, 1992; de Koning et al., 2000a). Increased levels and pulsatility of FSH in the follicular phase are seen in mothers of hereditary dizygotic twins (Lambalk et al., 1998). FSH receptor variants have now been identified (Perez et al., 2000; Sudo et al., 2002). This may cause slightly diminished receptor function, resulting in higher levels of FSH to achieve adequate receptor response whilst not necessarily affecting ovarian reserve (Bukulmez and Arici, 2004).

The sensitivity of bFSH in identifying women who will not become pregnant with IVF has been calculated to be only 8%, whereas it is considered to be a specific test, in that 98% of women who achieved pregnancy had normal findings (Barnhart and Osheroff, 1998). The positive predictive value (the probability of not becoming pregnant given a positive test result) of the test is difficult to interpret, especially in younger patients or the general subfertile population where the prevalence of non-pregnancy is lower. As such there is a danger of overinterpreting its predictive value (Barnhart and Osheroff, 1999).

The reasons for these conflicting reports are not entirely clear, but are certainly related to the varying methodologies employed by different investigators, especially with regard to different threshold values for FSH, study groups, outcome measures and their definitions, duration of follow-up and data analysis (Table II).

Efforts to increase the sensitivity of prediction have been attempted by combining FSH with LH as a ratio of FSH:LH, with limited success (Mukherjee et al., 1996; Kim et al., 1997).

Despite these limitations, more studies have been performed, and hence much more is known, about the predictive value of FSH than any other marker of ovarian reserve. This, combined with the relative practicality, patient tolerability and low cost of performing the test, is likely to result in bFSH remaining one of the most commonly performed tests of ovarian reserve for some time to come.

### Estradiol

The condensed follicular phase length in older women may be as a result of a more advanced follicular recruitment by cycle day 3. This early dominant follicle selection is expressed by high serum estradiol (E2) concentrations (Licciardi et al., 1995).

It has been shown in an assisted reproduction treatment population (where GnRH analogues were not administered) that increasing day 3 estradiol concentrations are associated with decreasing oocyte numbers and pregnancy rates (Licciardi et al., 1995), a correlation which has been repeated elsewhere (Smoorich et al., 1995). In patients with normal FSH levels, basal estradiol (bE2) has been shown to predict high cancellation rates and low oocyte yield in IVF (Evers et al., 1998). In another study, cancellation rates did correlate with bE2 levels but did not correlate with pregnancy outcome in those patients who were not cancelled (Frattarelli et al., 2000). Pregnancy rates have also been shown to be higher in a group of women undergoing in vitro maturation (Mikkelsen et al., 2001).

The predictive ability of bE2 is improved in patients of advanced reproductive age, especially when combined with bFSH (Buyalos et al., 1997). However, these observations have not been confirmed by others (Lee et al., 1988; Scott et al., 1989) (Table III).

No data are currently available regarding basal estradiol levels in the general subfertile population (Bukman and Heineman, 2001). Further studies including data on day 3 estradiol values and fecundity in spontaneous cycles are required before evaluating this parameter further.

### Inhibin B

Inhibins and activins are glycoproteins that belong to the transforming growth factor β (TGFβ) family. Activins are dimers of β subunits and act as functional antagonists of inhibin to stimulate pituitary FSH synthesis and secretion (Muttukrishna and Knight, 1991). Inhibins are heterodimers consisting of two dissimilar subunits (α and β) linked by disulphide bridges, and are secreted by the ovarian granulosa and luteal cells during the menstrual cycle (Lockwood et al., 1998).

Inhibins are a part of the HPO axis, specifically having an inhibitory effect on pituitary FSH synthesis and secretion (Muttukrishna and Knight, 1990) (Figure 2). Serum dimeric inhibin B is regarded as a direct measure of ovarian reserve as it is mainly secreted by pre-antral follicles (Klein et al., 1996b), whereas inhibin A is produced primarily during the late
Inhibin B rises from early in the follicular phase to reach a peak coincident with the onset of the mid-follicular phase decline in FSH concentrations, it then declines during the luteal phase apart from a peri-ovular peak, which may represent release of follicular inhibin B from the rupturing follicle into the circulation (Muttukrishna et al., 1994; Groome et al., 1996). Inhibin B has also been shown to respond to exogenous FSH administration in the follicular phase (Burger et al., 1998). The close temporal relationship between changes in levels of inhibin B and FSH in the mid-follicular phase suggest that the release of inhibin B by the pre-ovulatory follicle critically regulates pituitary FSH secretion, with inhibin B exhibiting a distinct periodicity in normal women (Lockwood et al., 1998). Studies in older women reveal significantly lower levels of inhibin B in women aged 40–50 years with raised FSH compared to women with normal FSH (Klein et al., 1996b; Muttukrishna et al., 2000). Additionally, concentrations of inhibin A and B have been shown to be lower in women with elevated FSH and regular menstrual cycles (de Koning et al., 2000b).

Studies in an assisted reproduction population demonstrated that women with a low cycle day 3 inhibin B concentration (<45 g/ml) had a poorer response to ovulation induction and decreased likelihood of achieving pregnancy compared with women who had higher day 3 inhibin B levels (Seifer et al., 1997). Decreased inhibin B was also found in women with normal FSH levels implying that decreased inhibin B precedes a rise in FSH (Seifer et al., 1999). Doubt has been expressed, however, regarding the quality of the inhibin B assays used in these earlier studies (Bancsi et al., 1997).

Correlation with other markers of ovarian reserve has been shown (Tinkanen et al., 2001). There is evidence that early follicular inhibin B levels correlate with follicle cohort size (Elting et al., 2001) and oocytes retrieved following controlled ovarian stimulation with FSH (Eldar-Geva et al., 2000, 2002), and with an improved sensitivity and specificity compared to other basal markers of ovarian reserve (Ficicioglu et al., 2003). Another study which also showed good correlation with oocytes retrieved

Inhibin A and B production from the granulosa cells of the ovarian follicle at different stages of development.

Fertility in female cancer survivors

Table III. Basal estradiol as a predictor of fertility

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Population Study design</td>
<td>ART Prospective</td>
<td>ART Prospective</td>
<td>ART Prospective</td>
<td>ART Prospective cohort</td>
<td>ART Retrospective</td>
</tr>
<tr>
<td>Patients ± cycles</td>
<td>452</td>
<td>225/292</td>
<td>119/380</td>
<td>231</td>
<td>2476</td>
</tr>
<tr>
<td>Upper threshold (pg/ml)</td>
<td>75.0</td>
<td>80.0</td>
<td>80.0</td>
<td>60.0</td>
<td>80.0</td>
</tr>
<tr>
<td>Outcome measure</td>
<td>PR PR, CR</td>
<td>Cumulative CPR, LBR</td>
<td>PT age, RO, FR, CPR</td>
<td>bE2 predicts CR</td>
<td>bE2 predicts CR</td>
</tr>
<tr>
<td>Conclusion</td>
<td>Ongoing PR decreases with increasing bE2 can predict high CR and low PR independent of bFSH</td>
<td>bE2 can predict pregnancy outcome better than age and bFSH</td>
<td>bE2 can predict response to IVF in presence of normal bFSH</td>
<td>bE2 predicts CR but not pregnancy outcome if ongoing pregnancy</td>
<td></td>
</tr>
</tbody>
</table>

ART = assisted reproduction treatment; PR = pregnancy rate; CR = cancellation rate; CPR = clinical pregnancy rate; FR = fertilisation rate; LBR = live birth rate; RO = retrieved oocytes; bE2 = basal estradiol; bFSH = basal FSH.

Figure 2. Inhibin A and B production from the granulosa cells of the ovarian follicle at different stages of development.

Figure 3. The Hypothalamic–Pituitary–Ovarian (HPO) axis and the physiologic basis for biochemical tests of ovarian reserve. The ovary is shown during different phases of follicular maturation. Components of dynamic ovarian reserve testing are shown at the points on the HPO axis where they act. +VE = positive feedback, −VE = negative feedback.
K.Lutchman Singh, M.Davies and R.Chatterjee

Table IV. Inhibin B as a basal determinant of ovarian reserve

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>ART</td>
<td>ART</td>
<td>ART</td>
<td>ART</td>
<td>ART</td>
</tr>
<tr>
<td>Study design</td>
<td>Retrospective</td>
<td>Prospective</td>
<td>Retrospective</td>
<td>Prospective</td>
<td>Prospective RCT</td>
</tr>
<tr>
<td>Patients ± cycles</td>
<td>156/178</td>
<td>106</td>
<td>80</td>
<td>141.0</td>
<td>58</td>
</tr>
<tr>
<td>Threshold (pg/ml)</td>
<td>45.0</td>
<td>45.0</td>
<td>141.0</td>
<td>N/A (increment)</td>
<td>N/A (increment)</td>
</tr>
<tr>
<td>Outcome measure</td>
<td>RO, PR</td>
<td>PR</td>
<td>RO, PR, CR</td>
<td>RO</td>
<td>RO</td>
</tr>
<tr>
<td>Conclusion</td>
<td>Low inhibin B (&lt;45.0 pg/ml) predicts RO and PR</td>
<td>Day 3 and day 10 inhibin B does not correlate with PR</td>
<td>Day 5 inhibin B predicts CR</td>
<td>Basal inhibin B of limited value in predicting RO</td>
<td>Inhibin B increment (in EFORT) correlates best with RO</td>
</tr>
</tbody>
</table>

ART = assisted reproduction treatment; RO = retrieved oocytes; PR = pregnancy rate; CR = cancellation rate; EFORT = exogenous FSH ovarian reserve test.

Inhibin B levels on cycle day 5 have been shown to be predictive as an early indicator of response during ovarian stimulation as well as outcome (Penarrubia et al., 2000; Fawzy et al., 2002).

The predictive ability of basal inhibin B with regard to pregnancy has not been confirmed by several authors, leading to doubt regarding the clinical usefulness of day 3 inhibin levels in clinical practice (Corson et al., 1999; Hall et al., 1999; Creus et al., 2000; Dumesic et al., 2001). Studies regarding basal inhibin A have also shown a link with ovarian function and follicular development, but limited value in predicting IVF outcome (Hall et al., 1999; Casper et al., 2001) (Table IV).

Inhibin B has been evaluated in predicting response as a component of dynamic testing (Ficicioglu et al., 2003). These include the clomiphene citrate challenge test (CCCT) (Hofmann et al., 1998), GnRH agonist test (GAST) (Ravhon et al., 2000) and the exogenous FSH stimulation test (EFORT) (Dzik et al., 2000; Eldar-Geva et al., 2002). These studies are all limited by the fact that although they show correlation with ovarian response, predictive value was not assessed (Welt, 2002). More recently, however, studies suggest that the incremental increase in inhibin B following ovarian stimulation may be more predictive of assisted reproduction treatment outcome than basal levels of inhibin B. In a prospective study, the inhibin B response to exogenous FSH correlated well with the antral follicle count (AFC) and the number of oocytes received (Yong et al., 2003). Kwee et al. (2003) undertook a prospective randomized study, which compared endocrine markers of ovarian reserve using the CCCT and EFORT to predict ovarian response, and revealed a high correlation with inhibin B in EFORT for number of follicles after ovarian stimulation, with increments in estradiol and inhibin B providing prediction of ovarian capacity.

The reasons for these conflicting reports are not clear and may reflect inter-assay variability and differing reference points and standards between different institutions, resulting in reduced specificity of the inhibin assay. Given the expense of performing the test, further prospective evaluation and confirmation of its predictive capacity is required before inhibin B can assume a routine place in the clinical assessment of women with reduced ovarian reserve. Until then, its measurement in the field of reproductive medicine should only be as part of a research protocol (Tong et al., 2003).

Anti-Müllerian hormone (AMH)

In females, the granulosa cells of the ovary produce AMH, which is a member of the transforming growth factor β (TGFβ) family. Levels can be almost undetectable at birth (Rajpert-De Meyts et al., 1999), with a subtle increase noted after puberty (Hudson et al., 1990). Serum levels on day 3 of the menstrual cycle show a progressive decrease with age, which correlates with antral follicle counts (de Vet et al., 2002).

The exact physiological basis of AMH is poorly understood in humans. The main physiological effect is the regression of Müllerian ducts in male fetuses (Josso et al., 1998). Based mainly on in vitro work and experiments performed in rats, AMH seems to play a pivotal role with regard to follicle recruitment, in that it inhibits the recruitment of primordial follicles into the pool of growing follicles and also decreases the responsiveness of growing follicles to FSH (Durlinger et al., 2002). There is evidence that AMH is mainly expressed in pre-antral and early antral follicles (Baarends et al., 1995). AMH levels are also seen to decline gradually during multiple follicular maturation as part of controlled ovarian stimulation (COS). This may be accounted for by the increased number of mature follicles expressing less AMH (Fanchin et al., 2003a). These findings suggest that AMH is solely produced by antral follicles capable of growing, and as such serum levels of AMH may represent both the quantity and quality of the ovarian follicle pool (te Velde and Pearson, 2002).

Recent studies have shown an association between reduced baseline serum AMH and poor response to IVF (van Rooij et al., 2002), and that initial AMH is associated with ovarian response in IVF patients who have normal FSH levels (Seifer et al., 2002). The relationship between AMH and other markers of ovarian reserve has been assessed, where AMH was found to correlate more strongly with the antral follicle count (AFC) than basal inhibin B, E2, FSH and LH in an infertile population (Gruijters et al., 2003). On this basis further validation of the test seems justified.
Normative data have been published (Cook et al., 2000; Laven et al., 2004) but are lacking with regard to day-to-day fluctuations, pulsatility and inter- or intra-cycle variability. The assay is expensive, and although several sensitive assays have already been developed (Long et al., 2000), the problem of inter-assay variability and the setting up of standardized reference values must be addressed. More research is required into the relationship between AMH and ovarian follicle dynamics, and as a marker of ovarian reserve in the general population before allowing its entry into clinical practice.

Dynamic ovarian reserve tests

The recognized limitations of biochemical markers of ovarian reserve led to the development of a number of dynamic ovarian tests in an attempt to ‘unmask’ patients whose diminished ovarian reserve led to the development of a number of dynamic ovarian reserve tests. The GAST/G-test: This test evaluates the estradiol serum concentration change before allowing its entry into clinical practice.

GnRH agonist stimulation test (GAST/G-test):

This test evaluates the estradiol serum concentration change from cycle day 2 to day 3 after administration of a GnRH agonist, the latter causing a temporary increase in pituitary secretion of FSH and LH (Sharara et al., 1998).

Early studies were able to show a limited ability for the early rise in estradiol concentration to differentiate between normal and reduced ovarian reserve, albeit with different threshold values (Padilla et al., 1990). Others have looked at the sum of stimulated and basal FSH concentrations with some success, but was this not superior to bFSH (Galtier-Dereure et al., 1996). In a significant study, the actual increase in estradiol concentration from day 2 to day 3 (ΔE2) following GnRH agonist administration on day 2 was a better predictor of ovarian reserve than age, bFSH and FSH:LH ratio (Ranieri et al., 1998). The same authors were also able to demonstrate that such an assessment would also allow effective drug regimen selection for IVF (Ranieri et al., 2001).

Because inhibin B is considered a direct marker of ovarian reserve, combining it with the GAST may have improved the predictive value of the test for IVF outcome. In fact, changes in the concentrations of inhibin B and estradiol have both been shown to correlate highly with the ovarian response to stimulation for IVF treatment (Ravhon et al., 2000) (Table VI).

The GAST appears to have much promise as a dynamic test of ovarian reserve but has not yet been validated outside an assisted reproductive setting. Furthermore the cost of performing the test has limited its acceptability in the general infertility population.

Clomiphene citrate challenge test (CCCT)

First described by Navot et al. (1987), this test involves the administration of 100 mg clomiphene citrate on cycle days 5–9, and the determination of FSH concentrations on days 3 and 10. In women with normal ovarian reserve, the overall increase in estradiol and inhibin production by the developing follicles should be able to overcome the estrogen antagonist effect of clomiphene on the HPO axis, and suppress FSH levels back into the normal range by day 10.

The evidence for this physiological suppressive effect of inhibin was supported by a study which showed reduced inhibin B levels following the CCCT in women with reduced ovarian reserve (Hofmann et al., 1998). In this way, the CCCT may be more indicative of oocyte quality than quantity. Nevertheless, a quantitative relationship between the CCCT and follicle density has been shown (Gulekli et al., 1999).

Table VI. GnRH agonist stimulation test in assisted reproduction

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>ART</td>
<td>ART</td>
<td>ART</td>
<td>ART</td>
</tr>
<tr>
<td>Study design</td>
<td>Comparative</td>
<td>Prospective</td>
<td>Prospective</td>
<td>Prospective</td>
</tr>
<tr>
<td>Patients ± cycles</td>
<td>100</td>
<td>107</td>
<td>119</td>
<td>75</td>
</tr>
<tr>
<td>Threshold (pg/ml)</td>
<td>Day 2:day 5 ratio = 2</td>
<td>ΔE2 100</td>
<td>RO</td>
<td>RO/FSH dose</td>
</tr>
<tr>
<td>Outcome measure</td>
<td>PR</td>
<td>PR</td>
<td>ΔE2 predicts RO</td>
<td>ΔE2 correlates with ovarian response</td>
</tr>
<tr>
<td>Conclusion</td>
<td>E2 response predicts PR</td>
<td>ΔE2 predicts PR</td>
<td>ΔE2 predicts RO</td>
<td>ΔE2 correlates with ovarian response</td>
</tr>
</tbody>
</table>

ART = assisted reproduction treatment; PR = pregnancy; rate RO = retrieved oocytes; ΔE2 = delta (stimulated–baseline) estradiol.
Several studies have now been published showing that women with a normal test respond better to ovarian stimulation, whereas an excessive response predicted a poor outcome to COS for IVF (Loumaye et al., 1992; Tanbo et al., 1992; Hofmann et al., 1996; Csemiczky et al., 2000). Correlation with biophysical markers of ovarian reserve (ovarian volume and AFC) has been shown in an infertile population (Erdem et al., 2004). The CCCT also predicts for improved pregnancy rates compared to those of ovarian reserve (Kwee et al., 2003). The patient's age should not be ignored, however, as pregnancy rates still diminish in patients with advancing age despite a normal CCCT (Scott et al., 2003). Further prospective evaluation is required to validate this test properly.

**Exogenous FSH ovarian reserve test (EFORT)**

This test combines bFSH with the rise in estradiol ($\Delta E_2$) over a 24 h period after administration of a standardized dose (300 IU) of purified FSH on day3.

Using this test, one group was able to improve the predictive value of bFSH alone for IVF outcome in stimulated cycles (Fanchin et al., 1994). A recent randomized, prospective study in which basal and dynamic tests of ovarian reserve were compared revealed the EFORT as being the best predictor of ovarian reserve (Kwee et al., 2003). Further prospective evaluation is required to validate this test properly.

**hMG stimulation test**

In the most recent study using this test, basal values of FSH, $E_2$ and inhibin were compared with hormonal and ultrasound parameters performed after 5 days stimulation with hMG (Fabregues et al., 2000). Although the predictive value of the rise in $E_2$ for response to IVF was better than bFSH alone, it was not very specific. Furthermore, the ability to predict pregnancy was less than that provided by the woman's age (Fabregues et al., 2000). These results and the expense of performing it may explain the relative lack of popularity in performing this dynamic test in routine practice (Sharif et al., 1998) (Table VIII).

### Table VIII. Exogenous FSH ovarian reserve test (EFORT) and hMG stimulation tests in IVF

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>ART</td>
<td>ART</td>
<td>ART</td>
</tr>
<tr>
<td>Study design</td>
<td>Prospective</td>
<td>Prospective RCT</td>
<td></td>
</tr>
<tr>
<td>Patients ± cycles</td>
<td>52</td>
<td>80</td>
<td>110</td>
</tr>
<tr>
<td>Threshold</td>
<td>$\Delta E_2 &gt; 30$ pg/ml; bFSH &lt; 11 IU/l</td>
<td>NS</td>
<td>N/A (increment)</td>
</tr>
<tr>
<td>Outcome measure</td>
<td>Ongoing PR, CR</td>
<td>PR, CR</td>
<td>RO</td>
</tr>
<tr>
<td>Conclusion</td>
<td>$\Delta E_2$ superior to bFSH in predicting outcome</td>
<td>$\Delta E_2$ predicts CR but not superior to age in predicting PR</td>
<td>Inhibin B increment in EFORT best predictor of RO</td>
</tr>
</tbody>
</table>

### Notes:

- ART = assisted reproduction treatment; PR = pregnancy rate; CR = cancellation rate; RO = retrieved oocytes; NS = not specified; N/A = not applicable; bFSH = basal FSH.
Biophysical ORT

Because ovarian morphology is examined, these tests offer an opportunity to assess follicle dynamics directly and as such are considered direct measures of ovarian reserve. One could deduce that the number of remaining follicles in the ovary is related to fertility or the likelihood of entering menopause (Faddy et al., 1992). The routine presence of transvaginal ultrasound in general fertility and gynaecology clinics may also contribute to the increased popularity of using these methods.

Ovarian volume

The human ovary is a dynamic organ, which changes in size and activity throughout its lifetime. Mean ovarian volume increases from 0.7 cm³ at age 10 years to 5.8 cm³ at age 17 years (Ivarsson et al., 1983). Although there are limited data on ovarian volume in normal fertile women (Lass and Brinsden, 1999), studies suggest that there are no major changes in ovarian volume during the reproductive period in individual women until the pre-menopausal period (Christensen et al., 1997; Ng et al., 2003). Ovarian size has been shown to decrease in women aged >40 years, a trend which is apparently unaffected by parity (Andolf et al., 1987). After the menopause the ovaries shrink with a sharp fall in ovarian volume, a decrease which progresses with each decade of life (Pavlak et al., 2000). These findings have led to the evaluation of ovarian volume as a useful marker of menopausal status (Flaws et al., 2000, 2001). There is also evidence that ovarian volume may be an earlier indicator of post-menopausal status than menstrual status (te Velde et al., 1998; Burger, 1999).

With regard to reproductive age, an early study found a correlation between ovarian volume and reproductive outcome in IVF cycles, but no correlation was found with chronological age (Syrop et al., 1995). Stronger correlations have since been shown between ovarian volume and COS (Lass et al., 1997b; Sharara and McClamrock, 1999). Small ovarian volume (<3 cm³) is associated with a poor response to hMG and a very high cancellation rate during IVF (Syrop et al., 1999). One study was able to show a relationship between ovarian volume and the number of follicles before stimulation, but not the number of oocytes retrieved (Tomas et al., 1997), a finding that has been confirmed elsewhere (Tinkanen et al., 1999; Dunesci et al., 2001). When likelihood ratios are calculated from some of these studies, the overall conclusion is that ovarian volume is not a good predictor of pregnancy (Bukman and Heineman, 2001), but when a threshold value of 3 cm³ is used, ovarian volume can be predictive of failure of follicular stimulation (Lass and Brinsden, 1999).

Wallace and Kelsey (2004) have recently proposed a model using ovarian volume to predict reproductive age. By estimating the mean primordial follicle population in women aged 25–51 years using the Faddy–Gosden equation (Faddy and Gosden, 1996), and correlating it with mean ovarian volume at each chronological age as described by Pavlik et al. (2000), reproductive age (age at menopause in this model) can be estimated (Wallace and Kelsey, 2004). This model makes a critical assumption that primordial follicle populations are fixed at birth, a notion that has recently been challenged (Johnson et al., 2004a). Whether or not ovarian volume can accurately reflect primordial follicle numbers or indeed follicle quality is questionable. The latter may be more important if reproductive age is to be correlated with fertility outcome.

Ovarian volume estimation using conventional ultrasound is inexpensive and relatively easy to perform. Intra- and inter-observer variations are small, ensuring reproducibility (Higgins et al., 1990; Lass et al., 1997b). Its main limitation at present is the lack of data regarding ovarian volume measurements in the general population, both fertile and infertile, across different age cohorts. When this is achieved, ovarian volume measurements are likely to play an important role in the clinical estimation of ovarian reserve (Table IX).

Antral follicle count (AFC)

Since the relationship between declining AFC and increasing age was first described (Ruess et al., 1996), many attempts have been made to establish a relationship between AFC and ovarian reserve. The assumption made is that the number of antral follicles originating from the cohort of growing follicles also correlates with the number of primordial follicles, or ovarian reserve (Scheffer et al., 1999). AFC is usually defined as the number of follicles <10 mm in diameter detected by ultrasound in the early follicular phase.

Studies in women with proven fertility reveal a continual decline of AFC with age, and with AFC showing better correlation with chronological age than other hormonal and ultrasound markers. (Ng et al., 2003). AFC has been shown to be predictive of ovarian response in IVF (Tomas et al., 1997; Chang et al., 1998a,b; Ng et al., 2000; Pohl et al., 2000; Hsieh et al., 2001; Ng et al., 2003).

Table IX. Ovarian volume as a predictor reproductive age

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>ART</td>
<td>ART</td>
<td>ART</td>
<td>ART</td>
<td>General</td>
</tr>
<tr>
<td>Study design</td>
<td>Retrospective</td>
<td>Prospective</td>
<td>Prospective</td>
<td>Comparative</td>
<td>Cross-sectional</td>
</tr>
<tr>
<td>Patients ± cycles</td>
<td>188</td>
<td>140</td>
<td>109</td>
<td>261</td>
<td>50</td>
</tr>
<tr>
<td>Threshold/cm³</td>
<td>Smallest OV 3 cm³</td>
<td>MOV 3 cm³</td>
<td>MOV 3 cm³</td>
<td>Smallest OV 5 cm³</td>
<td>Variable</td>
</tr>
<tr>
<td>Outcome measure</td>
<td>PR, CR</td>
<td>PR, CR</td>
<td>CR</td>
<td>RO, CR, PR</td>
<td>Menopausal status</td>
</tr>
<tr>
<td>Conclusion</td>
<td>OV (total and smallest) predicts response</td>
<td>Low MOV correlates with CR</td>
<td>Low MOV correlates with CR</td>
<td>OV superior to bFSH and bE₂</td>
<td>OV has similar sensitivity and specificity to bFSH and age status</td>
</tr>
</tbody>
</table>

ART = assisted reproduction treatment; OV = ovarian volume; MOV = mean ovarian volume; PR = pregnancy rate; CR = cancellation rate; RO = retrieved oocytes.
Nahum et al., 2001; Frattarelli et al., 2003). Compared with other markers of ovarian reserve, AFC was the best predictor of response to exogenous gonadotrophins (Bancsi et al., 2002) and correlates independently with the number of oocytes retrieved during IVF (Yong et al., 2003).

Reproducibility has been addressed and has been described as being moderate, with decreasing agreement at higher AFC (Scheffer et al., 2002); however, this study was limited to two observers. Inter-cycle variability has also been described as moderate in a fertile population (Scheffer et al., 1999). A more recent analysis, which studied infertile patients, showed that intercycle variability was greater than inter-observer variations, suggesting a biological variation of AFC from cycle to cycle (Hansen et al., 2003). As has been reported by other investigators, identification of an optimal cycle for stimulation did not predict for significantly better outcomes, and as such a single measurement of AFC is recommended (Hansen et al., 2003; Bancsi et al., 2004) (Table X).

Overall the results of these studies show considerable promise for the use of AFC in the estimation of ovarian reserve. Like ovarian volume, the procedure itself is relatively inexpensive and easy to perform, and can provide information immediately before biochemical results are obtainable. More data are required in different cohorts and a more rigorous assessment of inter-observer variability is required before it becomes part of routine clinical practice.

Ovarian stromal blood flow (OSBF)

The rationale for this test lies in the supposition that the primordial follicles in the ovary have no independent capillary network, and therefore depend on their proximity to the stromal vessels for their supply of nutrients and hormones (Findlay, 1986).

A relationship has been established between OSBF velocity and ovarian follicular response (Zaidi et al., 1996). This was confirmed in a cohort of patients with normal bFSH undergoing IVF treatment following pituitary suppression, in which mean ovarian stromal peak systolic velocity (PSV) was a better predictor of ovarian responsiveness than age (Engmann et al., 1999). In another study based on a fertile population, however, age was not related to mean PSV (Ng et al., 2003). Studies have shown a relationship between peri-follicular blood flow Doppler indices and oocyte developmental competence/implantation potential (Nargund et al., 1996; Chui et al., 1997; Huey et al., 1999) and others have reported that the mean flow index (intensity) after pituitary suppression correlates with the number of oocytes retrieved (Kupesic and Kurjak, 2002).

A limitation of this technique is that it is very operator dependent. It would require an ultrasonographer of considerable experience to obtain accurately the angle between the ultrasound beam and the intra-ovarian vessels in order to measure peak systolic velocity. This may result in a prolonged scanning time for acquisition of data. It has been suggested that 3D ultrasound can circumvent this difficulty (Kupesic and Kurjak, 2002). Using vascularization flow indices, 3D power Doppler ultrasound has been used to show that flow intensity in the ovarian stroma decreases with age (Pan et al., 2002). Another study was unable to show a correlation between these indices and ovarian response during IVF (Jarvela et al., 2003a), with inter- and intra-observer variability being described as acceptable (Jarvela et al., 2003b).

Further study is required to assess the predictive value and reproducibility of ovarian stromal blood flow measurements compared to other biophysical markers of ovarian reserve before entering routine clinical practice.

Three-dimensional ultrasound

It has been suggested that 3D ultrasound can improve the prediction of ovarian response to gonadotrophin stimulation, compared to conventional transvaginal ultrasound with power Doppler imaging. Three-dimensional imaging yields a more accurate estimate of ovarian volume, endometrial volume AFC and assessment of ovarian stromal perfusion by 3D power Doppler (Kupesic, 2001). Initial attempts to evaluate the use of 3D ultrasound to evaluate low responders to stimulation for IVF proved that the technique was useful in differentiating poor responders who had normal baseline FSH levels (Pellicer et al., 1998). This study did not attempt to compare 3D ultrasound with conventional transvaginal ultrasound imaging, however. Another study evaluated the use of AFC, ovarian volume, stromal area and OSBF as assessed by 3D ultrasound in predicting ovarian response and IVF outcome (Kupesic and Kurjak, 2002). In this study, OSBF and AFC seemed to be better predictors than ovarian volume and total stromal area (Kupesic and Kurjak, 2002). The same investigators have also been able to show recently that increasing patient age is associated with poor ovarian response,
as represented by reduced ovarian volume, AFC and poor stromal vascularity as assessed by 3D ultrasound (Kupesic et al., 2003) (Table XI).

One significant advantage of 3D ultrasound over conventional transvaginal ultrasound that has emerged from these studies is the reduction in examination time required for these patients. Imaging data can be stored allowing further assessment to be performed without the presence of the patient. A significant disadvantage is the high cost of 3D ultrasound over conventional ultrasound. More studies are required to compare the predictive ability of both modalities prospectively. Until these data become available, 3D ultrasound would not be expected to replace conventional ultrasound for some time to come.

**Ovarian follicle density (OFD)**

The histopathological examination of ovarian biopsy specimens has been advocated as a direct marker of ovarian reserve. Assessment of follicular density from ovarian biopsies taken from 60 subfertile women with normal bFSH levels showed a decline with increasing age, but did not follow up for pregnancy outcome (Lass et al., 1997a). A study performed on a cohort of patients with Turner’s syndrome who were undergoing ovarian tissue cryopreservation showed that OFD correlated with levels of FSH (Hreinsson et al., 2002). In a unique study, women aged >35 years undergoing oophorectomy for uterine pathology had indirect markers of ovarian reserve assessed in a previous cycle (Gulekli et al., 1999). Although a positive correlation was identified between basal serum E2 concentrations and follicle density in ovarian tissue, no significant correlation was seen between basal and clomiphene-stimulated FSH levels and OFD (Gulekli et al., 1999).

The main problem with this technique lies in the possibility that the follicle density measured in a particular specimen (biopsy site or histological section) may not reflect the picture in the rest of the ovary (Lass, 2001). This was shown recently, where the follicle density of ovarian biopsy specimens obtained at the time of ovarian tissue cryopreservation varied by more than two orders of magnitude, although there was a significant inverse linear correlation with age (Schmidt et al., 2003). Lambalk et al. (2004) attempted to investigate this further by performing multiple biopsies of 2 and 5 mm and whole ovaries from five patients of reproductive age, during operations not involving ovarian pathology. The authors attempted to estimate total ovary follicle number by transforming biopsy counts based on surface area. Despite this, predictive values based on the biopsies were widely varied (Lambalk et al., 2004). At present, there are no standardized measures for OFD among histopathology laboratories. This ‘correction factor’ is necessary to account for that proportion of the ovary not included in the sampling analysis, and may account for the large discrepancy in follicle numbers reported by various laboratories (Tilly, 2003).

Another drawback pertains to acquisition of an adequate specimen, which includes the invasiveness of performing a biopsy and the risk of adhesion formation (Sharara and Scott, 2004). One cannot put a patient through a surgical procedure to obtain a biopsy of ovarian tissue of unproven value.

In the specific situation of cancer patients undergoing ovarian tissue freezing for fertility preservation, OFD may be a useful research tool, as has been done elsewhere (Poiriot et al., 2002). This possibility, as well as progress in estimating the distribution of follicles in the ovarian cortex, still provides some hope for the clinical application of this test in the future (Lass, 2004; Sharara and Scott, 2004).

**Response to COS**

Although this group of patients is distinct from the general subfertile population, studies in this area have provided useful insight into the concept of ovarian reserve. One might consider it to be an extended form of dynamic testing.

Distinct cohorts of infertile women, with regular cycles (de Boer et al., 2003) and normal bFSH levels, who do not respond to COS, have been shown to develop POF (Farhi et al., 1997; Nikolaou et al., 2002; Lawson et al., 2003). This potential relationship has been challenged by others (De Sutter and Dhont, 2003). These women also seem to exhibit characteristics of reduced ovarian reserve when they are tested following a poor response to COS (Beckers et al., 2002). Despite these findings, women with a poor response to COS in the first IVF cycle can still have a normal response in subsequent cycles (Klinkert et al., 2002). What adds to the difficulty of interpreting these events is the lack of a uniform definition for poor response among different institutions (Lashen et al., 1999).

Response to COS has been attracting considerable interest recently as a marker of ovarian ageing. Although the model it

---

**Table XI. Three-dimensional ultrasound as a predictor of outcome in assisted reproduction treatment (ART)**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>ART</td>
<td>ART</td>
<td>ART</td>
</tr>
<tr>
<td>Study design</td>
<td>Prospective, case–control</td>
<td>Prospective</td>
<td>Prospective</td>
</tr>
<tr>
<td>Parameters measured</td>
<td>AFC and OV</td>
<td>AFC, OV, total stromal area OSBF</td>
<td>AFC, OV, OSBF</td>
</tr>
<tr>
<td>Patients ± cycle number</td>
<td>18</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>Outcome measure</td>
<td>Comparison between normal and low responders</td>
<td>RO, PR</td>
<td>RO, FR, PR</td>
</tr>
<tr>
<td>Conclusion</td>
<td>AFC but not OV discriminates normal from low responders</td>
<td>AFC and OSBF predicts outcome</td>
<td>All parameters can predict outcome regardless of age</td>
</tr>
</tbody>
</table>

ART = assisted reproduction treatment; AFC = antral follicle count; RO = retrieved oocytes; PR = pregnancy rate; FR = fertilization rate; OV = ovarian volume; OSBF = ovarian stromal blood flow.
provides is attractive, in that it involves a dynamic test with strictly controlled conditions and outcomes, it is limited by the fact that application to the general population is not appropriate.

ORT in cancer

The concept of ORT in patients with cancer is a relatively new one, and data pertaining to ovarian reserve tests in cancer are limited and mostly retrospective (Lower et al., 1999). The largest studies which attempted to predict the effects of adjuvant chemotherapy on ovarian function in pre-menopausal women with breast cancer have examined the incidence of amenorrhoea rates and premature menopause respectively (Bines et al., 1996; Goodwin et al., 1999). Overall, there is a lack of pertinent data relating to ORT in patients with cancer.

More progress has been achieved with regard to estimating ovarian reserve in patients with haematological malignancy. For example, it has been shown in a prospective study that ovarian volume in patients with lymphoma and leukaemia treated by high dose chemotherapy and radiotherapy was reduced by 50% within 72 h of cessation of high dose therapy (Chatterjee et al., 1994). In another study using the hMG test there was diminished ovarian reserve in patients with lymphoma treated by BEAM chemotherapy, and many patients had diminished ovarian reserve even prior to chemotherapy due to antecedent chemotherapy or disease (Chatterjee and Goldstone, 1996). More recently, inhibin A and inhibin B (in addition to FSH, LH and estradiol), have been assessed prospectively as prognostic factors for resumption of ovarian function in patients with lymphoma who received GnRH analogue co-treatment, with encouraging results (Blumenfeld, 2002).

Role of ORT in cancer

Young women treated for cancer have many concerns and harbour questions about the effect their treatment will have on their fertility (Schover, 1999) and menopausal status (Muscari et al., 1999). Many patients are left with significant anxieties and insufficient information about reproductive issues (Schover et al., 1999). Recognition of the fact that successful treatment of childhood cancer can be associated with impaired gonadal function in adulthood (Thomson et al., 2002) has led to increased efforts to estimate ovarian reserve in long-term survivors of childhood cancer. Basal gonadotrophins are unreliable in pre-pubertal children due to the relative quiescence of the HPO axis. Some investigators have proposed the use of inhibin B levels where suppression may have been indicative of arrested follicle development (Crofton et al., 2003). Another study revealed that despite regular menstrual cycles, survivors of childhood cancer had smaller ovarian volume, reduced AFC and inhibin B than controls (Larsen et al., 2003b). The same authors showed that in a cancer survivor cohort with FSH levels <10 IU/L, ovarian reserve was still reduced as evidenced by smaller ovarian volumes and lower AFC compared with age-matched controls (Larsen et al., 2003a).

Despite these findings, the data cannot necessarily be extrapolated to patients with different types of cancer because of differences in disease, type of chemotherapy and the age group of study.

Figure 4. The potential role of ovarian reserve testing in patients with cancer.
A unique web-based survey of fertility issues in young women with breast cancer revealed that only 51% felt that their concerns were addressed adequately. Also, 29% of women reported that infertility concerns influenced treatment decisions (Partridge et al., 2004).

Estimation of functional ovarian reserve in these patients can have a significant impact on how patients are counselled both before and after chemotherapy. These tests can be applied both pre- and post-chemotherapy where patients can be classified into three groups: normal, poor and intermediate ovarian reserve. In the pre-chemotherapy patient, this information can be used to predict the response to ovarian stimulation for embryo/oocyte cryopreservation, allowing appropriate counselling and modification of treatment. For example, patients deemed to have normal ovarian reserve prior to potentially sterilizing chemotherapy might opt to have embryos cryopreserved knowing that their response to COS is likely to be good. The same may be applied to patients post-chemotherapy for treatment of subfertility. Alternatively, patients with normal ovarian reserve may be counselled regarding adequate contraception. Furthermore, the potential exists to counsel patients appropriately regarding the onset of a premature menopause.

In some cases, treatment decisions for the patient’s cancer can be affected. For example, in women with breast cancer where there is an option for either cytotoxic chemotherapy or reversible endocrine ablation, the latter may be preferred if ovarian reserve is already diminished prior to treatment (Figure 4).

One must stress, however, that these examples represent only the potential of ORT for these patients, as these applications remain as yet unproven. Furthermore, while fertility and quality of life issues are important, they must not be pursued at the expense of the patient’s overall treatment and welfare. While potentially informative, ORT can only help and clarify the decision-making process in a limited group of patients, as malignancy and reproductive issues are complex and difficult.

Summary of ORT

The recognized limitations of basal markers such as FSH have led to the evaluation of several other tests of ovarian reserve. Dynamic testing has the potential to unmask ‘poor responders’. Biophysical markers offer a more ‘direct’ estimate of ovarian reserve, and the widespread use of ultrasound scanning makes it a relatively easy application to integrate into practice.

Despite all the options available, the ideal ORT has yet to be identified. More evidence is required, using adequately controlled and prospective study designs outside a fertility setting.

Pre-menopausal women who have been exposed to gonadotoxic chemotherapy either in childhood or adulthood have specific gynaecological concerns. These include the prospect of subfertility and a premature menopause (and its effects), and methods of preserving fertility and ovarian function. As yet there is a paucity of data relating to the use of ORT in patients with cancer.

ORT can potentially provide a useful counselling tool and act as a guide to treatment, including fertility preservation. This can be achieved with adequate longitudinal, controlled data in this population, as data from the subfertile population may not be extrapolated. The aim would be to develop a clinical tool which may allow classification of ovarian reserve for example into normal, intermediate and poor. This tool would allow a more accurate prediction including future response to treatment. Finally, the objective of this information would always be for the benefit of the psychological and physical well-being of the patient.

Conclusions

- Survival trends for cancer in pre-menopausal women and children are steadily improving.
- These patients represent a distinct cohort with specific concerns regarding their fertility (achieving pregnancy or adequate contraception) and the possibility of having a premature menopause.
- ORT has become established in the fertility setting, where it is used as a counselling tool and a guide to treatment.
- The ideal ORT has yet to be identified.
- The tests, which appear to show significant promise, include baseline inhibin B and AMH, stimulated inhibin B and estradiol, as well as AFC and ovarian volume.
- There is a lack of pertinent data relating to ORT in cancer, and information pertaining to ORT in a subfertile population cannot be extrapolated to include patients with cancer, as the pathophysiology is distinct.
- Patients with cancer would benefit from knowledge of their functional ovarian reserve. With this information, ORT could be applied clinically in patients with cancer as a guide to counselling and treatment.

Acknowledgements

The authors would like to thank Dr Shanthi Muttukrishna for her assistance in developing Figures 2 and 3.

References

Bancsi LF, Broekmans FJ, Looman CW, Habbema JD and te Velde ER (2004) Impact of repeated antral follicle counts on the prediction of
K.Lutchman Singh, M.Davies and R.Chatterjee

Barnhart K and Osheroff J (1999) We are overinterpreting the predictive value of serum follicle-stimulating hormone levels. Fertil Steril 72,8–9.
Chatterjee R and Goldstone AH (1996) Gonadal damage and effects on fertility in adult patients with haematological malignancy undergoing stem cell transplantation. Bone Marrow Transplant 17,5–11.
Chuang CC, Chen CD, Chao KH, Chen SU, Ho HN and Yang YS (2003) Age is a better predictor of pregnancy potential than basal follicle-stimulating hormone levels in women undergoing in vitro fertilization. Fertil Steril 79,63–68.


Espinoza MA, Coutinho C and Barnhart KT (2002) A moderately elevated day 3 FSH concentration has limited predictive value, especially in younger women. Hum Reprod 17,118–123.


Fertility in female cancer survivors
Fertility in female cancer survivors


Seifer DB, Scott RT, Jr, Bergh PA, Abrogast LF, Friedman CI, Mack CK and Danforth DR (1999) Women with declining ovarian reserve may demonstrate a decrease in day 3 serum inhibin B before a rise in day 3 follicle-stimulating hormone. Fertil Steril 72,63–65.


Yong PY, Baird DT, Thong KJ, McNeilly AS and Anderson RA (2003) Prospective analysis of the relationships between the ovarian follicle cohort and basal FSH concentration, the inhibin response to exogenous FSH and ovarian follicle number at different stages of the normal menstrual cycle and after pituitary down-regulation. Hum Reprod 18, 35–44.


Submitted on September 5, 2004; revised on October 21, 2004; accepted on October 28, 2004.