Published in final edited form as:

Fertil Steril. 2008 December ; 90(6): 2196–2202. doi:10.1016/j.fertnstert.2007.10.080.

A new approach to ovarian reserve testing

Wenjie Sun, M.D.^a, Barbara J. Stegmann, M.D., M.P.H.^{b,c}, Melinda Henne, M.D., M.S.^c, William H. Catherino, M.D., Ph.D.^{c,d}, and James H. Segars, M.D.^{b,c,d}

^a Department of Obstetrics and Gynecology, Wake Forest University Baptist Medical Center, Medical Center Blvd. Winston-Salem, NC, 27157

^b Reproductive Biology and Medicine Branch, NICHD, 10 Center Dr., Bldg. 10, Rm. 1E-3140, Bethesda, MD 20892-1109

^c Walter Reed ART Program, Walter Reed Army Medical Center, 6900 Georgia Ave., NW, Washington, DC 20307

^d Uniformed Services University of Health Sciences, Dept. of Obstetrics and Gynecology Building A, Room 3077, 4301 Jones Bridge Road, Bethesda, MD 20814-4799

Abstract

Objective—To critically examine ovarian reserve testing prior to assisted reproduction.

Design—A PUBMED computer search to identify relevant literature.

Setting—Multiple sites

Interventions—Testing for ovarian reserve

Main Outcome Measures—IVF and pregnancy outcomes

Result(s)—The prevalence of ovarian insufficiency varies significantly for women aged 30–45 years. Generalization or averaging of threshold values across different aged women leads to very poor sensitivity, specificity, and positive predictive value for all tests of ovarian reserve. Because of the changing prevalence of ovarian insufficiency, there is no single, suitable threshold value for any screening test of ovarian reserve. Our analysis supports dividing impaired ovarian reserve into two groups: age-dependent ovarian aging (physiologic) and premature (non-physiologic) reductions in the oocyte pool. Interpretation of any screening test used requires that age is considered as a variable. To guide clinical interpretation of test results, we suggest using a nomogram of FSH values versus expected delivery rate-per-cycle-start with ART for a given age.

Conclusions(s)—Proper interpretation of screening tests for ovarian insufficiency in couples considering ART is important since the presence of impaired ovarian reserve is associated with a low likelihood of pregnancy. The condition of premature (non-physiologic) ovarian insufficiency warrants additional research.

Corresponding author: James H. Segars, M.D. RBMB, NICHD, National Institutes of Health, Bldg. 10, CRC, Room 1E-3140, MSC 1109, 10 Center Drive, Bethesda MD, 20892-1109, segarsj@mail.nih.gov.

Disclaimer: The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the department of the Army or the Department of Defense.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Ovarian reserve; ovarian insufficiency; FSH; CCCT; inhibin; DOR; diminished ovarian reserve; assisted reproduction; in vitro fertilization; IVF; Assisted Reproductive Technologies; ART

More than 100 years ago, population studies clearly documented a decrease in fertility with increasing age (1). In today's culture of widely available birth control and workforce equality, women often delay child-bearing to pursue a career. As a result, the childbearing age for women has been delayed from the 20's to the 30's and even into the early 40s (2). This societal shift has resulted in an increase in the number of women who are interested in fertility and have regular cycles, but who are subfertile due to a reduction in their oocyte (egg) supply. Recognition of the profound adverse effect of a reduction in oocyte supply on fertility led to the concept of ovarian reserve and the moniker of 'Diminished Ovarian Reserve' (DOR). The term was coined by Navot et al. in 1987 (3) for women having an "exaggerated follicle stimulating (FSH) hormone level of 26 IU/L or more (over 2 standard deviations above control value)" during a clomiphene citrate challenge test.

Women with diminished ovarian reserve have no overt clinical symptoms other than subfertility but do demonstrate subtle changes in baseline hormone levels. To help elucidate diagnosis of DOR, Navot et al. (3), Scott et al. (4), Toner et al. (5), and Hofmann et al. (6) developed clinical definitions that coupled abnormal ovarian reserve testing with a poor ovarian response. These reports documented higher cancellation rates and reduced live birth rates with assisted reproduction (3–6) in women with DOR. Later studies (7–11) also reported a clinical association between DOR and increased miscarriage rates. The condition of impaired ovarian reserve is not only a problem with patients undergoing ART, but is a recognized concern for patients undergoing all types of infertility treatments (12).

Scott et al. (12) estimated the overall prevalence of impaired ovarian reserve to be approximately 10% in a general infertility practice, with an increase in prevalence among older women. For example, of 236 infertile couples with an abnormal clomiphene citrate challenge test (CCCT), two of 61 (3%) occurred in patients younger than 30 years, five of 72 (7%) were seen in women aged 30–34, seven of 68 (10%) were in women aged 35–39, and nine of 35 (26%) occurred in women 40 or older (12). These results have been confirmed in a number of subsequent studies and a well-documented relationship exists between basal FSH and increasing age (5,13,14).

The pathophysiology associated with the age-related reduction in fertility (1) is a reduction in the oocyte pool accompanied by an increase in FSH due to reduced feedback from estrogen and inhibins (15). There are also reduced levels of hormones such as anti-Mullerian hormone (AMH) (16). Theoretically, in some women the reduction in the oocyte pool might be detected by a day 3 FSH or an AMH level, in others only by provocative testing such as a CCCT. The term 'diminished ovarian reserve' attempts to define a threshold value above which a reduction in the oocyte pool is associated with impaired fertility (4).

Screening for Ovarian Insufficiency

Many investigators have used a single test to predict ovarian response to stimulation and resulting pregnancy rates (4,5,15–19). Such tests include basal FSH on cycle day 3 (4,16), CCCT (17,20,21), inhibin-B (21–23), estradiol (21,24), anti-Mullerian hormone (25,26), antral follicle counts (AFC) (16,27) and ovarian volume (16,21). The success of each test can be measured against ovarian response and/or live birth rate per cycle. Many studies only reported ovarian response, not live birth rate per cycle. Ovarian response is an indirect measurement of the status of the primordial follicle pool. While the establishment of a pregnancy may be

influenced by the number of retrieved oocytes, delivery rate is the true outcome of interest, and therefore the important outcome to measure.

It is also important to note that there has been significant variability in testing and studies have used different criteria or thresholds to define 'poor ovarian response.' As a result, various sensitivities and specificities of the screening tests are reported. Table 1 (15,16,21,22,26–28) compares a selected summary of screening tests commonly used to evaluate for ovarian response in ART with reported sensitivities and specificities. Table 1 also includes the results of Esposito et al. (29) that correlated day 3 FSH values to live birth rate per ART cycle with its associated sensitivities and specificities. As suggested by Bancsi et al. (16) and Broekmans et al. (15), no single test clearly emerges that possesses a satisfactory sensitivity and specificity for clinical application. Even if live birth is measured (29), as the threshold considered to be normal is increased, there is improvement in specificity, but a corresponding reduction in sensitivity. This means that many patients will be stimulated without success because the sensitivity of the test is exceedingly low (about 10%).

Combinations of screening tests for Ovarian Insufficiency

Because single tests do not adequately identify women with impaired ovarian reserve, investigators have combined tests in an effort to distinguish normal women from those with impaired ovarian reserve (20). By performing sequential tests consisting of the mean of 2 cycles of day 3 FSH followed by a day 10 inhibin-B, Hendriks et al. (20) reported an increased probability of predicting impaired ovarian reserve. This combination of tests exhibited a sensitivity of 71%, a specificity of 98%, and an receiver operator characteristic curve (ROC) of 0.92, which is greater than the ROC for any single test (20). Similarly, Bancsi et al. (16) used a combination of day 3 FSH, basal inhibin-B levels and antral follicle counts to obtain a sensitivity of 75%, a specificity of 95% and ROC of 0.92. Sequential testing and combinations of tests were an improvement over the performance of a single test, but the reports did not account for the different prevalence of ovarian insufficiency in women of different ages.

The conundrum of any screening test for Ovarian Insufficiency

While sequential testing improves the likelihood of diagnosing impaired ovarian reserve when compared to single tests, this is not a satisfactory solution because of one simple fact: the prevalence of ovarian insufficiency changes with age and the performance of the test changes as the prevalence changes. This point is illustrated in Table 2. In this example, the prevalence of impaired ovarian reserve changes from 3% in women of 30 years to virtually 100% in women over age 45. When disease prevalence is low, the positive predictive value (PPV) of the test is low, regardless of the sensitivity or specificity of the test. Given a test with 95% sensitivity and specificity and an expected prevalence of disease of 3%, a test would have a PPV of only 27%; that is there is a 27% probability the woman truly has impaired ovarian reserve if she has a positive test. If the prevalence is increased to 10% (such as for a 38 year old woman), then the probability of having diminished ovarian reserve based on this test increases to 67%, but this degree of certainty is only slightly greater than chance, or 50:50. Conversely, for the 40-year-old, an elevated FSH is much more likely to reflect the presence disease as the positive predictive value increases to 89% when the prevalence increases to 40%.

Contrary to intuition, changing the sensitivity of the test does not correct the problem. In the previous example, as the sensitivity increases to 98%, the specificity falls to 60%; but more importantly the positive predictive value in the 30 year old remains low (5%) because the prevalence of the disease has not changed (Table 2). Of note, the published sensitivities of tests for ovarian reserve screening (Table 1) fall within the values used in this example. In older women, because of the higher prevalence of impaired ovarian reserve, screening tests are more

likely to accurately predict impaired ovarian response, even at levels of sensitivity and specificity observed with tests currently in clinical use.

A solution to the problem of Ovarian Insufficiency screening

To fully understand why all tests perform poorly, it is necessary to distinguish physiologic ovarian aging from premature ovarian insufficiency. As illustrated schematically in Figure 1A, a number of studies have shown that physiologic aging leads to an age-related decline in delivery rates per ART cycle start. Conceptually, optimal fertility may be considered as fertility observed before age 30. Based on this concept, physiologic, age-dependent impairment in ovarian reserve can be understood as a reduction from that maximal level (dashed line), as shown by the hatched area (region a, Fig. 1A). In contrast, non-physiologic impairment in ovarian reserve cases that are not explained by age may be conceptualized as the reduction in fertility in the shaded area (region b, Fig 1A).

Why draw this distinction? Most importantly, because distinguishing physiologic ovarian insufficiency due to age from non-physiologic ovarian insufficiency identifies a separate group of women. This distinction is necessary in order to identify women with premature impairment of ovarian function for additional counseling, perhaps involving altered stimulation regimens. For example, Abdalla and Thum (30) showed that an elevated FSH was associated with an increase in the amount of gonadotropins needed as well as a higher cancellation rate in all age groups, and recommended altering the counseling these women received to reflect these facts.

Drawing this distinction is also essential in order to clarify the relationship of screening tests for ovarian reserve to ART outcomes. *The changing prevalence of ovarian insufficiency due to age complicates all screening tests for ovarian reserve.* The changing prevalence of the impaired ovarian reserve causes considerable overlap in test results between the two groups of women: those with physiologic, age-related ovarian insufficiency and those with non-physiologic impaired ovarian reserve. This overlap results in poor screening test performance—particularly if as has been suggested that increased FSH values in younger women may be more reflective of quantity, rather than quality of oocytes (30).

This distinction naturally highlights a second problem regarding the current strategies for detection of impaired ovarian reserve: the use of a single threshold value. As shown above, there is no single threshold or ‘cut-off value’ to define impaired ovarian reserve, rather there is a smooth continuum in the relationship of the physiologic reduction in oocyte pool upon which non-physiologic reductions in the oocyte pool may, or may not, be superimposed. By unbundling non-physiologic and physiologic ovarian insufficiency, it may then be possible to define the possible contribution of each mechanism.

The changing prevalence of disease means for the woman in her early 30s, if a day 3 FSH is “normal”, impaired ovarian reserve *is not likely* to be present. If the 30-year-old has a positive test (e.g., an elevated day 3 FSH), then a false positive test is certainly possible and *additional assessment* is indicated. In contrast, day 3 FSH testing for the 38-year-old woman is more predictive of poor outcome. However, if the test is negative, the patient may still respond poorly, as the negative predictive value (NPV) of the test becomes less reliable as the prevalence of the disease increases. This is precisely what is observed clinically: screening tests of ovarian reserve are more predictive in older women.

Because there is no universally accepted threshold, opinions vary on how to define clinically significant reductions in ovarian reserve. Some ART clinics have chosen to abandon any screening and proceed to stimulate all candidates using a ‘proof-in-the-pudding’ type of logic (16). Of course, this practice leads to high cancellation rates, or oocyte retrieval for patients

for whom live birth rates are expected to be less than 5%, at great cost to patients (31). Therefore, a more cost-effective approach is certainly warranted.

Interpretation of screening tests for Ovarian Insufficiency

Rather than a single threshold value for women of all ages, an enlightened approach to the problem would be to interpret ovarian reserve tests based on an individual patient's age. Using results published by Akande (32), it is possible to define an acceptable likelihood of delivery per cycle start, for instance at 5% (dashed line, Figure 1B), and then interpret values of basal FSH testing in light of whatever delivery rate is acceptable for the provider and patient. As illustrated in Figure 1B, the threshold level of FSH changes as a function of patient age. Use of a nomogram such as this provides a realistic assessment of the likelihood for delivery (not just pregnancy). Based on this approach, at age 35 and below, *any test suggesting ovarian insufficiency would need to be confirmed with additional testing*, for example an antral follicle count.

This additional testing would either support or refute the diagnosis of non-physiologic ovarian insufficiency by increasing the pre-test probability of the condition and, as shown by Hendriks et al. (33), could improve the test performance. Clearly, the thresholds for women under age 35 will be very different from women over age 40, and these thresholds should be based on the nomogram. By age 38, however, ancillary tests are no longer required, since a positive day 3 FSH or CCCT has a relatively high predictive value. Although this concept remains to be proven through prospective clinical studies, such an approach would be expected to improve upon the poor predictive value of screening tests. It is worth noting that the overall delivery rates published by Akande et al. (32) are low in comparison to U.S. results, but the relationship of FSH and outcome persists even when delivery rates are higher.

Should screening tests for Ovarian Insufficiency be abandoned?

While interpretation of screening tests for impaired ovarian reserve may be confusing, appropriate screening and diagnosis is important. There are significant consequences associated with the diagnosis, such as reduction in pregnancy rate, increased miscarriage rate (6–11,34), and increased cost per delivery (31,35–37).

The increase in cost due to impaired ovarian reserve was illustrated in a study of 1238 first IVF cycles at Walter Reed Army Medical Center from January 1st, 1999 to December 31st, 2003 (31). In another study (37), calculations were based on cycle start to account for the actual cancellation rates in each age group and to give a more accurate estimate of the impact of ovarian insufficiency and spontaneous abortion rates on live birth. Using the average cost of a single cycle in 2005, \$9,674, as estimated by Collins (38) and adjusted by the National Consumer Price Index for medical care services, the cost of a live birth following ART in a woman age 35 with an FSH < 10 IU/L was \$28,698 and increased to \$36,579 when FSH was > 10 IU/L (37). At age 42, the cost for delivery with an FSH < 10 IU/L was \$122,178, compared to \$309,925 when the basal FSH was 10 or greater (37). Clearly, the cost of ART warrants development of improved screening and diagnostic testing.

Other potential benefits of separating age-related from non-physiologic Ovarian Insufficiency

Women with impaired ovarian reserve not only have reduced likelihood of pregnancy, but may also have a significantly increased miscarriage rates (6–11,34). This is felt by some to be due to poorer oocyte quality and/or increased fetal aneuploidy rates (39–44), although Abdalla and Thum were unable to confirm this association in their study (30).

The observed increase in aneuploidy raises the question: should women with non-physiologic ovarian insufficiency have different screening criteria for fetal chromosomal abnormality testing than the general population? Amniocentesis is currently recommended for women of advanced maternal age, or in the presence of abnormal ultrasound findings or an abnormal triple screen test, but not for women with non-physiologic ovarian insufficiency. As the incidence of aneuploidy is increased in women with ovarian insufficiency regardless of age (8), these women might also benefit from screening, a point that requires additional prospective study.

Conclusions

Ovarian insufficiency is an insidious process that begins years before the cessation of menstrual flow. While it is a physiologic consequence of aging, a number of women are also affected by non-physiologic reductions in oocyte pool, independent of age. Due to the marked effect of age, we suggest use of a nomogram rather than a “threshold” or a specific cutoff to diagnose impaired ovarian reserve. Accurate assessment of ovarian insufficiency allows the practitioner to tailor treatment to a woman’s individual needs.

Acknowledgements

This research was supported, in part, by the Reproductive Biology and Medicine Branch of NICHD, NIH.

The authors would like to acknowledge helpful discussions with Alicia Armstrong M.D. and Mark Payson, M.D., the expertise of Jacques Cohen, Ph.D., Aidita N. James, M.D., Sasha Hennessy, M.D., Donna Hoover, M.D. Darshana Naik, M.D., Frederick Larsen, M.D. and fellows in the Reproductive Endocrine Fellowship at NICHD-USUHS for their contributions and discussions regarding ovarian reserve testing.

References

1. Duncan, M. Fecundity, fertility, sterility and allied topics. New York: Gryphon Editions; 1895.
2. Diczfalussy E. Contraception and society. *Eur J Contracept Reprod Health Care* 2002;7:199–209. [PubMed: 12648291]
3. Navot D, Rosenwaks Z, Margalioth EJ. Prognostic assessment of female fecundity. *Lancet* 1987;2:645–7. [PubMed: 2887939]
4. Scott RT, Toner JP, Muasher SJ, Oehninger S, Robinson S, Rosenwaks Z. Follicle-stimulating hormone levels on cycle day 3 are predictive of in vitro fertilization outcome. *Fertil Steril* 1989;51:651–4. [PubMed: 2494082]
5. 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;55:784–91. [PubMed: 1901282]
6. Hofmann GE, Danforth DR, Seifer DB. Inhibin-B: the physiologic basis of the clomiphene citrate challenge test for ovarian reserve screening. *Fertil Steril* 1998;69:474–7. [PubMed: 9531880]
7. Trout SW, Seifer DB. Do women with unexplained recurrent pregnancy loss have higher day 3 serum FSH and estradiol values? *Fertil Steril* 2000;74:335–7. [PubMed: 10927054]
8. Levi AJ, Raynault MF, Bergh PA, Drews MR, Miller BT, Scott RT Jr. Reproductive outcome in patients with diminished ovarian reserve. *Fertil Steril* 2001;76:666–9. [PubMed: 11591396]
9. Kuo PL, Guo HR. Mechanism of recurrent spontaneous abortions in women with mosaicism of X-chromosome aneuploidies. *Fertil Steril* 2004;82:1594–601. [PubMed: 15589865]
10. Meden-Vrtovec H. Ovarian aging and infertility. *Clin Exp Obstet Gynecol* 2004;31:5–8. [PubMed: 14998176]
11. Elter K, Kavak ZN, Gokaslan H, Pekin T. Antral follicle assessment after down-regulation may be a useful tool for predicting pregnancy loss in in vitro fertilization pregnancies. *Gynecol Endocrinol* 2005;21:33–7. [PubMed: 16048799]

12. Scott RT, Leonardi MR, Hofmann GE, Illions EH, Neal GS, Navot D. A prospective evaluation of clomiphene citrate challenge test screening of the general infertility population. *Obstet Gynecol* 1993;82:539–44. [PubMed: 8377979]
13. Bruin, Jd; Velde, Et. Female reproductive ageing: concepts and consequences. In: Tulandi, T.; Gosden, R., editors. *Preservation of Fertility*. London, UK: Taylor & Francis; 2004. p. 3
14. Velde, Et; Pearson, P. The variability of female reproductive ageing. *Human Reproduction Update* 2002;8:141–54. [PubMed: 12099629]
15. 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;12:685–718. [PubMed: 16891297]
16. Bancsi LF, Broekmans FJ, Eijkemans MJ, de Jong FH, Habbema JD, te Velde ER. Predictors of poor ovarian response in in vitro fertilization: a prospective study comparing basal markers of ovarian reserve. *Fertil Steril* 2002;77:328–36. [PubMed: 11821092]
17. Csemiczky G, Harlin J, Fried G. Predictive power of clomiphene citrate challenge test for failure of in vitro fertilization treatment. *Acta Obstet Gynecol Scand* 2002;81:954–61. [PubMed: 12366487]
18. Sharif K, Elgendy M, Lashen H, Afnan M. Age and basal follicle stimulating hormone as predictors of in vitro fertilisation outcome. *Br J Obstet Gynaecol* 1998;105:107–12. [PubMed: 9442172]
19. Bancsi LF, Broekmans FJ, Mol BW, Habbema JD, te Velde ER. Performance of basal follicle-stimulating hormone in the prediction of poor ovarian response and failure to become pregnant after in vitro fertilization: a meta-analysis. *Fertil Steril* 2003;79:1091–100. [PubMed: 12738501]
20. Hendriks DJ, Broekmans FJ, Bancsi LF, de Jong FH, Looman CW, Te Velde ER. Repeated clomiphene citrate challenge testing in the prediction of outcome in IVF: a comparison with basal markers for ovarian reserve. *Hum Reprod* 2005;20:163–9. [PubMed: 15471926]
21. Erdem M, Erdem A, Gursay R, Biberoglu K. Comparison of basal and clomiphene citrate induced FSH and inhibin B, ovarian volume and antral follicle counts as ovarian reserve tests and predictors of poor ovarian response in IVF. *J Assist Reprod Genet* 2004;21:37–45. [PubMed: 15186020]
22. Balasch J, Creus M, Fabregues F, Carmona F, Casamitjana R, Ascaso C, et al. Inhibin, follicle-stimulating hormone, and age as predictors of ovarian response in in vitro fertilization cycles stimulated with gonadotropin-releasing hormone agonist-gonadotropin treatment. *Am J Obstet Gynecol* 1996;175:1226–30. [PubMed: 8942492]
23. Seifer DB, Scott RT Jr, Bergh PA, Abrogast LK, Friedman CI, Mack CK, et al. 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* 1999;72:63–5. [PubMed: 10428149]
24. Frattarelli JL, Bergh PA, Drews MR, Sharara FI, Scott RT. Evaluation of basal estradiol levels in assisted reproductive technology cycles. *Fertil Steril* 2000;74:518–24. [PubMed: 10973649]
25. van Rooij IA, Broekmans FJ, te Velde ER, Fauser BC, Bancsi LF, de Jong FH, et al. Serum anti-Mullerian hormone levels: a novel measure of ovarian reserve. *Hum Reprod* 2002;17:3065–71. [PubMed: 12456604]
26. Ficioglu C, Kutlu T, Baglam E, Bakacak Z. Early follicular antimullerian hormone as an indicator of ovarian reserve. *Fertil Steril* 2006;85:592–6. [PubMed: 16500324]
27. Frattarelli JL, Levi AJ, Miller BT, Segars JH. A prospective assessment of the predictive value of basal antral follicles in in vitro fertilization cycles. *Fertil Steril* 2003;80:350–5. [PubMed: 12909498]
28. Muttukrishna S, McGarrigle H, Wakim R, Khadum I, Ranieri DM, Serhal P. Antral follicle count, anti-mullerian hormone and inhibin B: predictors of ovarian response in assisted reproductive technology? *BJOG* 2005;112:1384–90. [PubMed: 16167941]
29. Esposito MA, Coutifaris C, Barnhart KT. A moderately elevated day 3 FSH concentration has limited predictive value, especially in younger women. *Hum Reprod* 2002;17:118–23. [PubMed: 11756373]
30. Abdalla H, Thum MY. An elevated basal FSH reflects a quantitative rather than qualitative decline of the ovarian reserve. *Hum Reprod* 2004;19:893–8. [PubMed: 15016786]
31. Henne MB, Stegmann BJ, Neithardt AB, Catherino WH, Armstrong AY, Kao TC, Segars JH. The combined effect of age and basal FSH on the cost of a livebirth at ART. *Fertil Steril*. 2007 in press
32. Akande VA, Fleming CF, Hunt LP, Keay SD, Jenkins JM. Biological versus chronological ageing of oocytes, distinguishable by raised FSH levels in relation to the success of IVF treatment. *Hum Reprod* 2002;17:2003–8. [PubMed: 12151428]

33. Hendriks DJ, Mol BW, Bancsi LF, Te Velde ER, Broekmans FJ. 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;83:291–301. [PubMed: 15705365]
34. Gurbuz B, Yalti S, Ficicioglu C, Ozden S, Yildirim G, Sayar C. Basal hormone levels in women with recurrent pregnancy loss. *Gynecol Endocrinol* 2003;17:317–21. [PubMed: 14503976]
35. Van Voorhis BJ, Syrop CH. Cost-effective treatment for the couple with infertility. *Clin Obstet Gynecol* 2000;43:958–73. [PubMed: 11100309]
36. Jain T, Missmer SA, Hornstein MD. Trends in embryo-transfer practice and in outcomes of the use of assisted reproductive technology in the United States. *N Engl J Med* 2004;350:1639–45. [PubMed: 15084696]
37. Stegmann, BJ.; Henne, MB.; Neithardt, AB.; Catherino, WH.; Kao, TC.; Segars, JH. Cost analysis of a successful ART cycle in women with an elevated day 3 FSH. American Society for Reproductive Medicine; New Orleans. 2006.
38. Collins J. Cost-effectiveness of in vitro fertilization. *Semin Reprod Med* 2001;19:279–89. [PubMed: 11679909]
39. Munne S, Cohen J, Sable D. Preimplantation genetic diagnosis for advanced maternal age and other indications. *Fertil Steril* 2002;78:234–6. [PubMed: 12137856]
40. Warburton D. Biological aging and the etiology of aneuploidy. *Cytogenet Genome Res* 2005;111:266–72. [PubMed: 16192704]
41. van Montfrans JM, Dorland M, Oosterhuis GJ, van Vugt JM, Rekers-Mombarg LT, Lambalk CB. Increased concentrations of follicle-stimulating hormone in mothers of children with Down's syndrome. *Lancet* 1999;353:1853–4. [PubMed: 10359418]
42. Freeman SB, Yang Q, Allran K, Taft LF, Sherman SL. Women with a reduced ovarian complement may have an increased risk for a child with Down syndrome. *Am J Hum Genet* 2000;66:1680–3. [PubMed: 10733467]
43. van Montfrans JM, van Hooff MH, Martens F, Lambalk CB. Basal FSH, estradiol and inhibin B concentrations in women with a previous Down's syndrome affected pregnancy. *Hum Reprod* 2002;17:44–7. [PubMed: 11756360]
44. Nasser A, Mukherjee T, Grifo JA, Noyes N, Krey L, Copperman AB. Elevated day 3 serum follicle stimulating hormone and/or estradiol may predict fetal aneuploidy. *Fertil Steril* 1999;71:715–8. [PubMed: 10202884]

Figure 1.

Combined effect of age and FSH upon expected delivery rate with assisted reproduction. (A) Delivery rate per cycle start versus age. Optimal success at ART is present at age 30, denoted by the dotted line. Physiologic reduction delivery rate due to aging and the associated ovarian insufficiency is indicated by the hatched area (a) in red; and can be conceptualized as a reduction from that dotted line. In contrast, non-physiologic ovarian insufficiency associated with increased basal FSH (or other markers of ovarian insufficiency) indicated marked by (b) in the blue area is also associated with a reduction in live birth rate. (B) Nomogram of delivery rate per cycle-start versus age including basal FSH after Akande et al. (32). The dotted line represents a 5% expected delivery rate per cycle start. Note that the 5% expected delivery rate crosses FSH values at different ages. This nomogram (or a similar nomogram specific for an ART program) can be used to plot patient values of FSH and age to assess the likelihood of live birth with assisted reproduction. Use of a nomogram may provide more accurate assessment of ART outcome than either age or basal FSH as singular values.

Table 1
Performance of ovarian insufficiency testing in the prediction of poor ovarian response and live birth

	Continuous data vs. Threshold Value	Sensitivity	Specificity	Positive Likelihood Ratio	Predictive ability(ROC Analysis)	Ref
FSH^a	Continuous	0.51–0.63	0.81	2.68	0.77	(15,21,22,28)
	>10mIU/mL	0.26–0.87	0.60–0.97	0.9–21.8	NA	
	>12mIU/mL	0.24	1.00	0.24	NA	
CCCT^b	Days 3/10 ≥10mIU/mL	0.69	0.88	5.75	0.81	(15,21)
	Day 10 >10	0.65	0.87	5.0	NA	
	Day 10>15	0.35	0.96	8.1	NA	
Inhibin B	Continuous	0.52–0.69	0.63–0.80	1.93–3.45	0.71	(15,21,22,28)
	40 pg/mL	0.87	0.49–0.64	1.7–2.42	NA	
	<45 pg/mL	0.33–0.53	0.79–0.95	1.57–10.6	NA	
	<53.8 pg/mL	0.39	0.94	6.5	NA	
AMH^c	Continuous	Not reported	Not reported	Not reported	0.92	(15,21,26,28)
	<0.1 ng/mL	0.49–0.76	0.88–0.94	4.08–12.6	NA	
	<0.2 ng/mL	0.54–0.87	0.64–0.90	2.45–5.7	NA	
	0.25 ng/mL	0.91	0.91	10.1	NA	
Ovarian Volume	<0.3 ng/mL	0.60	0.89	5.6	NA	
	Continuous	0.81	0.81	4.26	0.82	(21)
	<2.98 cm ³	0.08–0.75	0.81–0.94	1.30–3.95	NA	
	<7 cm ³	0.39–0.55	0.67–0.85	1.67–2.51	NA	
	<8.6 cm ³	0.61	0.73	2.23	NA	
AFC^d	Continuous	0.75	0.63	2.03	0.78–0.80	(15,21,26–28)
	<4 antrals	0.09–0.86	0.84–0.97	3.3–5.4	NA	
	≤4 antrals	0.30–0.89	0.39–0.96	1.45	NA	
	<6 antrals	0.36–0.81	0.77–0.89	1.57–7.4	NA	
FSH^a(live birth)	>10mIU/mL	0.19	0.91	0.8 ^e	0.28 ^e	(29)
	10–11.4mIU/mL	0.09	0.91	0.71 ^e	0.28 ^e	
	>11.4mIU/mL	0.11	1.00	1.00 ^e	0.28 ^e	

^aDay 3 Follicle Stimulating Hormone Level

^b Clomiphene Citrate Challenge Test, measuring FSH on day 3 (baseline) and day 10

^c Anti-Müllerian Hormone

^d Antral follicle count

^e Calculated using prevalence data reported by Esposito et al. (2002)

NA—Not Applicable, predictive ability better measured using sensitivity and specificity

Table 2
Changes in the predictive ability of screening tests based on changes in prevalence of disease

	Age	Pretest probability	Pre-test odds	Post-test Odds	Positive predictive value	Negative predictive value
Probability of a test with 95% sensitivity and specificity	30	0.02 (2/100)	0.02	0.39	0.27	0.99
	38	0.10 (10/100)	0.11	2.09	0.68	0.99
	40	0.30 (30/100)	0.43	8.14	0.89	0.98
Probability of a test with 98% sensitivity and 60% specificity	30	0.02 (2/100)	0.02	0.05	0.05	0.99
	38	0.10 (10/100)	0.11	0.27	0.21	0.99
	40	0.30 (30/100)	0.43	1.05	0.51	0.99
Probability of a test with 71% sensitivity and 98% * specificity	30	0.02 (2/100)	0.02	0.72	0.42	0.99
	38	0.10 (10/100)	0.11	3.6	0.80	0.97
	40	0.30 (30/100)	0.43	15.2	0.94	0.89

* Best combination test using mean day 3 FSH + mean day 10 inhibin B over 2 cycles.

A combination test of AFC, inhibin B, and basal FSH yields a 75% sensitivity and a 95% specificity, thus yield post-test probability of 23%, 60%, 82% for 30, 38, and 40 year-old women consecutively.