EDITORIAL COMMENTARY

Perspectives on the efficacy and indications for preimplantation genetic screening: where are we now?

Marc A. Fritz

Division of Reproductive Endocrinology and Fertility, Department of Obstetrics and Gynecology, University of North Carolina at Chapel Hill, CB#7570, 4001 Old Clinic, Chapel Hill, NC 27599-7570, USA

Correspondence address. E-mail: marc_fritz@med.unc.edu

Keywords: preimplantation genetic screening; aneuploidy

Given that the majority of failed implantations and early pregnancies reasonably may be attributed to embryo aneuploidy, the logical foundation for the use of preimplantation genetic screening (PGS) in cycles of in vitro fertilization (IVF) seems undeniably sound. If abnormal embryos having no implantation or advanced developmental potential can be identified accurately and excluded, and if only normal, euploid embryos are transferred, improved outcomes certainly should be expected, at least in women at a high risk of aneuploidy, if not in all women. Those at greatest risk of aneuploidy and therefore presumed most likely to benefit from PGS include women of advanced maternal age, women having recurrent unexplained miscarriages or implantation failure after transfer of good-quality embryos, and women with partners having severe male factor infertility.

Results of a recent randomized controlled trial evaluating the efficacy of preimplantation genetic screening (PGS) for the indication of advanced maternal age stimulated a vigorous and passionate debate, which continues (Mastenbroek et al., 2007). This issue of Human Reproduction contains four new contributions relevant to that debate. One presents a subgroup analysis of data from the study by Mastenbroek et al., performed to assess whether the effect of PGS on live birth rates in women of advanced maternal age differed with varying risks for aneuploidy (Twisk et al., 2008). A second reports the results of a new randomized controlled trial examining the impact of PGS on in vitro fertilization (IVF) outcomes in women of advanced maternal age (Hardarson et al., 2008). The third offers an opinion about the wisdom and ethics of performing additional trials of PGS in women of advanced maternal age (Mastenbroek et al., 2008). The last reports the results of a randomized controlled trial evaluating the efficacy of PGS for improving the live birth rates achieved with single embryo transfer in young women (Staessen et al., 2008). A brief review of previous PGS trials will help to place these contributions in proper context and clearer perspective.

The concept of PGS is so logical and attractive that its efficacy was generally assumed and accepted without serious question, once its technical feasibility had been demonstrated (Munné et al., 1993a,b). Enthusiasm for PGS was fueled by the results of observational studies comparing the outcomes of IVF with and without PGS and finding that PGS was associated with increased implantation rates (Gianaroli et al., 1999; Munné et al., 2003) and decreased miscarriage rates (Munné et al., 2006). The use of PGS has increased steadily, rapidly in recent years, despite the lack of any substantive evidence that PGS can deliver on its promise for achieving higher live birth rates (Twisk et al., 2006). Some predicted that PGS was destined to become a routine element of all IVF cycles (Verlinsky et al., 2004), and in some highly competitive regions in the USA, marketing even implied that programs which could not or chose not to offer PGS were substandard.

The first serious doubts about the efficacy of PGS emerged from the results of a randomized controlled trial that appeared in 2004. Staessen et al. (2004) compared implantation rates in women aged 37 years and older after a single cycle of IVF with and without PGS using fluorescence in situ hybridization (FISH) for seven chromosomes (X, Y, 13, 16, 18, 21 and 22). In the PGS group, two blastomeres were removed from all embryos having six cells or more. Up to three (age 37–39 years), or six blastocysts were transferred (age ≥40 years), and in the PGS group, only chromosomally normal embryos were transferred. Significantly fewer cycles progressed to transfer in the PGS group (54.7%) than in control group (85.8%). Implantation rates (17.1% in the PGS group versus 11.5% in the control group) were not significantly different. When later analyzed by others, on an intention-to-treat basis, the ongoing pregnancy rate in the PGS group was 11.0%, compared with 15.3% in the controls [relative risk 0.72, 95% confidence interval (CI) 0.43–1.21] (Mastenbroek et al., 2005). The overall pregnancy loss rates (including both preclinical and clinical abortions) also were not different (24.1% in the
PGS group versus 25.6% in the control group (Staessen et al., 2004). Critics contended that (i) the removal of two blastomeres seriously compromised the developmental potential of embryos transferred in the PGS group, (ii) the low implantation rate of blastocysts reflected poor laboratory performance and (iii) the higher implantation rate in the PGS group provided evidence that selection based on PGS actually compensated for the harm resulting from biopsy (Cohen and Munne, 2005).

In 2007, Mastenbroek et al. (2007) reported the results of a large, multicenter, randomized, double-blind trial comparing ongoing pregnancy rates in women aged 35–41 years who received up to three cycles of IVF with and without PGS using FISH for eight chromosomes (1, 13, 16, 17, 18, 21, X and Y), performed on a single blastomere. In the PGS group, the two chromosomally normal embryos having the best morphological characteristics were transferred, but if no chromosomally normal embryos with good morphological features were available, embryos with good morphology but ‘undetermined’ chromosomal status were transferred. Compared with controls, women assigned to PGS achieved significantly lower rates of ongoing pregnancy (25 versus 37%; rate ratio 0.69, 95% CI 0.51–0.93) and live birth (24 versus 35%; rate ratio 0.68, 95% CI 0.50–0.92); miscarriage rates did not differ significantly between the two groups. The authors concluded that PGS did not increase, but instead significantly reduced ongoing pregnancy and live birth rates after IVF in women of advanced maternal age (Mastenbroek et al., 2007). Some agreed (Collins, 2007); others clearly did not (Handyside and Thornhill, 2007; Munne et al., 2007a,b; Wilton, 2007; Simpson, 2008).

Critics emphasized that the effect of PGS is highly dependent on the use of appropriate methods and on patient selection, and contended that serious methodological flaws prevented any confident conclusions regarding the potential benefits of PGS for women of advanced maternal age. The relatively high diagnostic failure rate (20%) and the low implantation rate of ‘undetermined’ embryos (6%), compared with that in controls (14.7%) were interpreted as indicating a high ‘biopsy-damage rate’ which negated the benefit of PGS that otherwise may have been realized (Munne et al., 2007a; Simpson, 2008). Other perceived flaws were the failure to exclude common aneuploidies involving chromosomes 15 and 22, and the inclusion of cycles in which fewer than six to eight embryos were available for analysis (Munne et al., 2007a). Some viewed the trial as demonstrating only that, in inexperienced hands, PGS can be detrimental (Munne et al., 2007b; Simpson, 2008). Some argued that if cycles in which ‘undetermined’ embryos had been transferred were excluded from the analysis, the implantation rate in the PGS group (16.8%) was higher than that in controls (14.7%), and because the miscarriage rates in the two groups were similar, the live birth rate after PGS was equivalent to, if not higher than in controls (Wilton, 2007; Simpson, 2008). Others interpreted the results as demonstrating only the need for safer biopsy, more accurate and comprehensive chromosome analysis, and targeted application to high-risk groups (Handyside and Thornhill, 2007). Those most circumspect concluded that only additional well-designed and well-executed randomized trials could resolve persistent questions about the benefits of PGS (increased live birth rates and decreased miscarriage rates), in relation to its associated risks and costs, and about the safety of biopsy, test accuracy and diagnostic efficiency, and the appropriate indications for PGS (Harper et al., 2008; Yakin and Urman, 2008).

The study by Twisk et al. that appears in this issue was aimed at one of the perceived flaws in the trial conducted by Mastenbroek et al.—that the benefits of PGS were obscured in an unselected population having varying risks of aneuploidy. Twisk et al. (2008) performed a subgroup analysis of data from the original study, dividing the population into groups based on factors known to affect the risk of embryonic aneuploidy, and compared directly the effect of PGS within the various groups to determine whether any subgroup might benefit from PGS. Acknowledging the limitations and pitfalls of subgroup analyses, the authors used interaction tests to restrain inappropriate subgroup findings while still having the ability to detect interactive effects. No significant differential effect of PGS was found in groups based on maternal age (<38 and ≥38 years), previous miscarriages (none, one and two or more), semen quality (total motile count <1 or ≥1 million), the total amount of recombinant FSH administered during stimulation (subdivision based on quartiles) or the total number of top-quality embryos (none, one or two, three or more). Finding that live birth rates after PGS were lower than those without PGS in all groups, the authors conclude that PGS has no clinical benefit in women of advanced maternal age, regardless of their risk of embryonic aneuploidy, and challenge the a priori assumption that PGS will benefit most, those at highest risk of aneuploidy.

The analysis by Twisk et al. (2008) is informative, offers additional new insights into the results reported by Mastenbroek et al. and is provocative. However, a subgroup analysis addressing only one of the several perceived methodological flaws in the original study that critics viewed as neither valid nor generalizable cannot be expected to settle the question or to quell the debate. Hardarson et al. (2008) report the results of yet another randomized controlled trial examining the impact of PGS on IVF outcomes in women of advanced maternal age (≥38 years). Clinical pregnancy rates were compared after a single cycle of IVF with and without PGS using FISH for seven chromosomes (13, 16, 18, 21, 22, X and Y), performed after removal of a single blastomere (in 91% of cases). In both groups, one or two embryos were transferred, and in the PGS group, only chromosomally normal embryos were transferred. In the control group, embryo transfer was performed earlier, on Day 3 after oocyte retrieval. Secondary outcomes included pregnancy rate per transfer, and rates of implantation, spontaneous miscarriage and delivery. The study was terminated early because an interim analysis, conducted after 109 of 320 projected subjects were enrolled (34%), revealed a very low conditional power (10%) to show superiority for PGS. Compared with controls, women in the PGS group achieved significantly lower rates of clinical pregnancy (8.9 versus 24.5%; rate difference 15.6%, 95% CI 1.8–29.4; P = 0.039) and live birth (5.4 versus 18.9%; rate difference 13.5%, 95% CI 1.4–25.6; P = 0.039). There were no differences between groups for
any other secondary outcomes. The authors interpret their results as providing further evidence against the use of PGS for the indication of advanced maternal age.

The study by Hardarson et al. (2008) represents the fifth randomized controlled trial examining the impact of PGS on IVF outcomes in women of advanced maternal age, including two relatively small trials thus far reported only in abstract form (Stevens et al., 2004; Debrock et al., 2007); all have failed to demonstrate any benefit from PGS. Of the five, three have found no difference between PGS and control groups (Staessen et al., 2004; Stevens et al., 2004; Debrock et al., 2007), and two have observed poorer outcomes after PGS (Mastenbroek et al., 2007; Hardarson et al., 2008). Thus, the volume of evidence against the use of PGS to improve IVF outcomes in older women has grown. However, many of the same criticisms of the study by Mastenbroek et al. can be mounted again by those so inclined. Although the diagnostic failure rate (9%) was lower than that reported by Mastenbroek et al. (20%), it was still more than twice that cited as typical in ‘experienced laboratories’ (≤4%) (Munné et al., 2007a). Although not significantly different, the lower implantation rate of biopsied embryos (11.4%), compared with that for Day 3 transferred embryos in controls (18.9%), might again be interpreted as evidence of ‘biopsy damage’ (Munné et al., 2007a; Simpson, 2008). Although the FISH panel did include chromosome 22, it did not include chromosome 15 (Munné et al., 2007a), or 17, and although ‘undetermined’ embryos were not transferred (Wilton, 2007; Simpson, 2008), women having as few as two embryos available for biopsy were included (Munné et al., 2007a). Consequently, this study too may be dismissed by critics. However, as the number of trials and independent centers with similar results increases, the argument that outcomes merely reflect a lack of technical skill (embryo biopsy) and/or laboratory expertise (genetic diagnosis) (Simpson, 2008) becomes increasingly difficult to sustain.

The newest contribution to the ongoing debate in this journal (What next for preimplantation genetic screening?), from Mastenbroek et al., questions the need for and the ethics of performing more well-designed and well-executed randomized clinical trials, as three earlier contributors to the debate had suggested (Harper et al., 2008; Jansen et al., 2008; Yakin and Urman, 2008). The authors performed a meta-analysis of randomized comparative data on PGS for the indication of advanced maternal age, including five trials involving a total of 7334 randomized cycles (Staessen et al., 2004; Stevens et al., 2004; Debrock et al., 2007; Mastenbroek et al., 2007; Hardarson et al., 2008). After finding that overall ongoing pregnancies were significantly reduced after PGS [odds ratio (OR) 0.56, 95% CI 0.42–0.76], they calculated the necessary power of a new trial to shift the common OR for ongoing pregnancy rate to significantly greater than 1.0, indicating a benefit from PGS, and determined that the trial would need to include at least 6000 cycles. In their view, it is unethical to perform such a trial, because participants may be exposed to an ineffective or harmful treatment, and because the intrinsic limitations of PGS, as it is currently performed, make it highly unlikely that the trial would yield evidence for benefit.

The opinion offered by Mastenbroek et al. (2008) is well-reasoned and merits careful consideration before the time and substantial resources required are devoted to the effort of a new and larger trial of PGS for the indication of advanced maternal age. Mastenbroek et al. calculated that at least 6000 cycles would be needed to shift the overall weight of evidence to benefit, with ongoing clinical pregnancy as the primary outcome. However, almost all would agree that the best outcome measure is live birth rate per started cycle, and a trial powered for that primary outcome would need to be larger still. Intermediate or surrogate end-points are informative, but frequently inconclusive. Measures wherein the denominator is the number of embryos transferred (implantation rate), the number of retrievals or transfers, or the number of pregnancies (pregnancy loss rate) may even be misleading because they do not account for the patients who never get that far. Moreover, live birth per started cycle is the outcome most useful in treatment planning, and arguably the outcome of greatest importance and interest to patients. Those who view all of the trials thus far as fatally flawed and who believe firmly in the value of PGS might argue that a new trial clearly is needed and would not need to be so large if designed to address effectively all of the real or perceived weaknesses of earlier trials. That may not be possible, but the results might again be challenged or dismissed if it does not.

Staessen et al. (2008) report the results of a randomized controlled trial to evaluate the impact of PGS on IVF outcomes achieved after single embryo transfer in young women (age <36 years), reasoning that PGS might aid in the selection of the embryo having the greatest implantation potential. Live birth rates per cycle start were compared after a single cycle of IVF with and without PGS using FISH for seven chromosomes (13, 16, 18, 21, 22, X and Y), performed after removal of one (85% of embryos) or two blastomeres (15% of embryos). By design, a single blastocyst was transferred on Day 5 after oocyte retrieval, and in the PGS group, only chromosomally normal embryos were transferred. Secondary outcomes included biochemical and clinical pregnancy rates, and miscarriage rates. The study was terminated prematurely at time of the first interim analysis, after 120 of 447 projected subjects (27%) in each group were enrolled. An ad hoc power analysis indicated that there was less than a 3% probability of achieving the defined objective of a significant 10% increase or decrease in live birth rate with PGS. Live birth rates per cycle start in the PGS and control groups were identical (34.6%). Further analysis revealed no significant influence of age, duration of infertility, parity, diagnosis, oocyte yield or sperm parameters on outcomes. Overall miscarriage rates in the two groups were not different; in the PGS group, all (n = 10) were preclinical losses, whereas in the control groups 7 of 15 were clinical losses. The authors conclude that PGS does not increase the delivery rate in young women receiving a single blastocyst transfer.

The results reported by Staessen et al. (2008) are not altogether surprising since, at least in theory, PGS should benefit most those at increased risk for aneuploidy, and young women unselected for past implantation failures or pregnancy losses generally are not. However, the purpose of
PGS as applied in the study was different than in previous trials—to maximize live birth rates achieved with single blastocyst transfer. That PGS failed to increase live birth rates over those achieved with blastocyst transfer alone is disappointing, because there was reason to hope that the strategy might increase the chance for a successful pregnancy while virtually eliminating the risk of a multiple birth. It is unclear whether PGS failed because (i) extended culture alone excludes the majority of abnormal embryos, (ii) biopsy negates any benefits that otherwise might derive from screening or (iii) embryonic mosaicism confounds and may even subvert the selection process by favoring selection of an embryo that screens normal, but is not, over a morphologically superior embryo that screens abnormal, but is not. From a clinical perspective, the reason also is moot. Staessen et al. found that although PGS did not improve outcomes, aneuploidy screening also had no apparent adverse effects. That finding contrasts with the results of another recent trial of similar design in ‘good prognosis’ patients under age 39 (mean age 31 years) that was terminated early for harm after observing significantly lower implantation rates in PGS cycles; later analysis also revealed a significantly lower live birth rate in the PGS group (Meyer et al., 2008). Another study aimed at improving outcomes achieved with single blastocyst transfer through PGS performed on trophoectoderm was suspended when interim analysis revealed a trend to lower success rates in the PGS group, and terminated early when subsequent comparisons demonstrated that live birth rates were significantly lower in the PGS group than in controls (Jansen et al., 2008). In the only other comparable trial, involving an unselected population of patients, results of an interim analysis revealed no significant benefits from PGS, although there was a trend toward improved pregnancy and live birth rates in PGS cycles (Mersereau et al., 2007). So, where are we now? Five trials examining the impact of PGS on outcomes in women of advanced maternal age, and four in patients having a generally good prognosis, all have failed to demonstrate any clear benefit for PGS. Moreover, there are no data from randomized controlled trials relating to the efficacy of PGS for improving live birth rates in patients with other proposed indications for PGS. In 2007, the American Society for Reproductive Medicine issued a Practice Committee opinion concluding that available evidence did not support the use of PGS as currently performed to improve live birth rates in patients with advanced maternal age, previous implantation failure or recurrent pregnancy loss or to reduce miscarriage rates in patients with recurrent pregnancy loss related to aneuploidy (Practice Committee of the Society for Assisted Reproductive Technology and Practice Committee of the American Society for Reproductive Medicine, 2007). A year later, those conclusions seem even more justified, and ‘patients who receive a single embryo transfer’ might now be added to the list.

PGS should work, but after more than a decade of experience, there is arguably no truly substantive evidence to indicate that it does. Several reasons have been suggested to explain why, all relating to the intrinsic limitations of PGS as currently performed, including (i) adverse effects of embryo biopsy on implantation or developmental potential, (ii) transfer of presumed normal embryos that were aneuploid for one or more chromosomes not analyzed and (iii) misdiagnoses due to errors in interpretation of mosaicism, resulting in inadvertent transfer of abnormal embryos or exclusion of normal embryos. Although it certainly is true that technical expertise increases with experience, the failure of PGS in randomized trials cannot reasonably be attributed solely to a lack of technical prowess. Nonetheless, if it is, and if PGS can be effectively applied by only a few, we must conclude that its utility is quite limited. Emerging technologies allowing for the analysis of all chromosomes, such as comparative genomic hybridization (Voulaire et al., 2000; Wilton, 2002), can overcome one of the current limitations of PGS, but the vexing problem of mosaicism will remain to be solved. New techniques and strategies for PGS will, and should, be sought and tested. However, experience has taught us they must be evaluated rigorously and must have demonstrable potential for benefit before they are applied widely, and prematurely.

References


