DEBATE—CONTINUED

What next for preimplantation genetic screening? More randomized controlled trials needed?

S. Mastenbroek1,6, P. Scriven2, M. Twisk1, S. Viville3,4,5, F. Van der Veen1 and S. Repping1

1Center for Reproductive Medicine, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands; 2Center for Preimplantation Genetic Diagnosis, Department of Cytogenetics, Guy’s and St Thomas’ NHS Foundation Trust, London, UK; 3Service de Biologie de la Reproduction—Syndicat Inter-Hospitalier de la Communauté Urbaine de Strasbourg, Centre Médico-Chirurgical et Obstétrical, Service de Biologie, Schiltigheim, France; 4Department of Developmental Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch, France; 5Faculté de Médecine, Centre Hospitalier Universitaire, Université Louis Pasteur, Strasbourg, France

6Correspondence address. Tel: +31-20-5663090; E-mail: s.mastenbroek@amc.uva.nl

The recent debate on preimplantation genetic screening (PGS) has raised questions about its routine use in clinical practice. It has been suggested that the most effective way to resolve the debate about the usefulness of PGS is to perform more well-designed and well-executed randomized controlled trials (RCTs). However, in view of the lack of evidence for the effectiveness of PGS and the accumulating evidence for its harmfulness, it is our opinion that it is unethical to perform additional RCTs for the indication advanced maternal age using cleavage stage biopsy.

Keywords: PGS; ethics; clinical practice

Preimplantation genetic screening (PGS) was introduced in clinical practice with the aim to improve pregnancy rates in subfertile couples, based on the assumption that high rates of chromosomal aneuploidy, frequently found in cleavage stage embryos of these couples, were responsible for the disappointingly low pregnancy rates after ART (Wilton, 2002). Since the first reported pregnancies after PGS in 1995 (Verlinsky et al., 1995), there has been a steady increase in the use of this technique. The most extensive registry available to date is that of the European Society for Human Reproduction and Embryology (ESHRE) PGD Consortium, which reported on 116 cycles of PGS performed worldwide in 1997–1998 and 2087 cycles in 2004 (Harper et al., 2008b). The data from 2005 follows this trend with 2275 cycles (Goossens et al., 2008). The total number of cycles performed worldwide each year is much higher, since only a limited number of PGS centers report their data to the consortium. A survey in 2005, among all US-based infertility centres, showed that 127 out of 186 centres (68%) performed PGS, with a total of 2197 cycles (Baruch et al., 2008).

In the ongoing debate series ‘what next for preimplantation genetic screening?’, it was concluded that there is no evidence of increased live birth rates after PGS and the suggestion was made to perform more well-designed and well-executed randomized controlled trials (RCTs) (Jansen et al., 2008; Yakin and Urman, 2008; Harper et al., 2008a).

In this contribution to the debate series, we wish to question the wisdom of this suggestion by a critical appraisal of the available data. To do so, we first performed a meta-analysis of randomized comparative data on PGS for the indication advanced maternal age that was published in peer-reviewed journals or in abstracts of scientific meetings. This meta-analysis shows a clear-cut and statistically significant reduction of ongoing pregnancies after PGS [odds ratio (OR): 0.56, 95% confidence interval (CI): 0.42–0.76] (Fig. 1). We then calculated the necessary power of a new RCT to shift the current common OR from 0.56 to a common OR significantly >1.0, indicative of a positive effect of PGS on ongoing pregnancy rates. Assuming a clinically relevant relative 20% increase in ongoing pregnancy rate after PGS in this new RCT, this trial should then include at least 6000 cycles and would yield a new common OR of 1.13 (95% CI: 1.01–1.26) (Fig. 2).

There are two reasons why it is unethical to perform such a trial. The first reason is that the most important ethical condition to perform any trial, i.e. the concept of being in equipoise, is not met (Lilford and Jackson, 1995; Lilford, 2003). It is, therefore, unethical to ask consent from potential donors and embryos if the outcome of an RCT is not expected to alter the current practice significantly.
participants and expose them to an ineffective, or even harmful, treatment. A well-known example of continuing with RCTs, when in fact a consistent statistically significant OR was already established, is the use of streptokinase as thrombolytic therapy for acute myocardial infarction, where in retrospect 30,000 patients were included in randomized trials over a period of 15 years only for narrowing CIs around the same mean effect in cumulative meta-analyses (Lau et al., 1992).

If the same is true for PGS, then a new RCT of 6000 women will do nothing more than narrowing CIs at the cost of reducing the number of ongoing pregnancies in the PGS arm of the trial from an expected 621 to 385. This reduction in ongoing pregnancies is estimated by multiplying the number of pregnancies that would have been in the PGS arm of the trial if PGS was not applied (621), i.e., the number of cycles in the PGS arm of the trial (3000) multiplied by the ongoing pregnancy rate in the control group (0.207), i.e., one minus the relative risk (calculated from the common OR 0.56 using ongoing pregnancy (0.207) in the control group according to Zhang et al. (1998)) of ongoing pregnancy after PGS (1–0.62).

The second reason is that there are intrinsic limitations of the current PGS techniques that make it highly unlikely that a new trial for PGS will ever find an increase in ongoing pregnancies. First, technical limitations apply to biopsy, fixation and FISH analysis, the three essential steps in PGS, which are all not without failure, not even in the hands of highly experienced personnel. Second, it is well-known that FISH analysis has in general a 92–99% accuracy per probe, so when using a multi-probe panel on one blastomere, the risk of misdiagnosis is significant. The low positive predictive value of the test will result in the exclusion of embryos for consideration for transfer that have the potential to be successful (Michiels et al., 2006; Deugarte et al., 2008). Third, embryo mosaicism, the condition that the chromosomal constitution differs between blastomeres of the same embryo, is a frequent phenomenon among human preimplantation embryos. As a consequence, the cell biopsied during PGS is in many cases not representative of the genotype of the embryo. In the case of diploid–aneuploid mosaicism, the most frequent form of mosaicism among human preimplantation embryos (Bielanska et al., 2002; Baart et al., 2006), aspiration of a normal blastomere will reduce the proportion of diploid blastomeres in the embryo and lead to transfer or cryopreservation of an embryo with an increased proportion of abnormal cells. Conversely, aspiration of an aneuploid blastomere will increase the proportion of normal blastomeres and lead to the discarding of these embryos, despite the fact that they have potential to be viable.

Figure 1: The effect of cleavage stage PGS in patients with advanced maternal age on the ongoing pregnancy rate per cycle. Results are presented as two types of meta-analyses. On the left is the traditional meta-analysis and on the right the same data are presented as cumulative meta-analyses (adapted from Antman et al., 1992).

<table>
<thead>
<tr>
<th>Study</th>
<th>Cycles</th>
<th>OR (fixed)</th>
<th>OR (fixed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>95% CI</td>
<td>95% CI</td>
</tr>
<tr>
<td>1 - Staessen 2004</td>
<td>289</td>
<td>0.67 [0.37, 1.24]</td>
<td></td>
</tr>
<tr>
<td>2 - Stevens 2004</td>
<td>39</td>
<td>0.42 [0.11, 1.62]</td>
<td></td>
</tr>
<tr>
<td>3 - Mastenbroek 2007</td>
<td>836</td>
<td>0.60 [0.41, 0.89]</td>
<td></td>
</tr>
<tr>
<td>4 - Debrock 2007</td>
<td>61</td>
<td>0.24 [0.09, 1.46]</td>
<td></td>
</tr>
<tr>
<td>5 - Hardarson 2008</td>
<td>109</td>
<td>0.24 [0.09, 1.46]</td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>1334</td>
<td>0.56 [0.42, 0.76]</td>
<td></td>
</tr>
</tbody>
</table>

Test for overall effect: Z = 3.75 (P = 0.0002)

Figure 2: Similar to Fig. 1, but now a hypothetical RCT was added to show the number of patients needed in a new trial before meta-analysis will show a significant beneficial effect of PGS. A 20% ongoing pregnancy rate in the control group was assumed for this hypothetical RCT (which is the mean of the ongoing pregnancy rates per cycle in the control groups of the previous five RCTs) and a relative increase of 20% in ongoing pregnancy rate per cycle after PGS was assumed. Such a hypothetical RCT will have to include at least 6000 cycles before meta-analysis will show a significant positive effect after PGS, despite the very unlikely assumed success rate (20%) after PGS.
The current available evidence on the efficacy of PGS has been put aside by some reasoning that the RCTs lack technical prowess causing them to be neither valid nor generalizable (Simpson, 2008). When five trials from five independent established groups all show the same negative effect of PGS, setting them all aside, because they were of insufficient quality, is too simple. In fact, since all RCTs show the same effect of PGS, it seems only justified to generalize the outcome of these studies and to conclude that there is no beneficial effect of PGS in terms of increased ongoing pregnancy rates. Furthermore, the intrinsic limitations of PGS as described above provide a much better explanation for the inefficacy of PGS.

Advanced maternal age is the most common indication for PGS, but PGS has also been applied for other indications like recurrent implantation failure, recurrent early pregnancy loss, severe male factor infertility and more recently for good prognosis patients (Harper et al., 2008b). Only one trial, presented as a poster at the ESHRE meeting in Barcelona, 2008 (Blockeel, P-522), investigated patients with repeated implantation failure. The trial included 140 cycles and showed a relative risk of clinical pregnancy per cycle after PGS of 0.60 (95% CI: 0.35 – 1.03). For the indications recurrent early pregnancy loss and severe male factor infertility, no trials have been performed. For good prognosis patients, i.e. younger infertile women (<38 years of age) with multiple good quality embryos, four trials with a total of 374 patients have been performed (Meyer et al., 2006; Mersereau et al., 2007; Staessen et al., 2008; Jansen et al., 2008). The evidence in these trials was insufficient to show benefit or harm.

When embarking on new RCTs for indications other than advanced maternal age using current techniques or new RCTs for any indication using emerging techniques, such as the analysis of all chromosomes, it is essential that pilot-studies should first clearly demonstrate the potential for benefit in terms of increased ongoing pregnancy rates per cycle. The lack of high-level evidence for the effectiveness of PGS after more than 10 years of practice and many thousands of cycles, the accumulating high-level evidence for no benefit, and even harm for women of advanced maternal age on one hand, and the unavoidable technical drawbacks and the obscuring effects of chromosomal mosaicism on the other hand, should make any other approach untenable.

References


Submitted on July 8, 2008; resubmitted on August 19, 2008; accepted on September 1, 2008.

Submitted on July 8, 2008; resubmitted on August 19, 2008; accepted on September 1, 2008.