Preimplantation genetic screening: a reappraisal
Mastenbroek, S.

Citation for published version (APA):
Chapter 6

Preimplantation genetic screening: a systematic review and meta-analysis of randomized controlled trials

Sebastiaan Mastenbroek
Moniek Twisk
Fulco van der Veen
Sjoerd Repping

Human Reproduction Update 2011; 17:454-66
Abstract

Background
Preimplantation genetic screening (PGS) has increasingly been used in the past decade. Here we present a systematic review and meta-analysis of RCTs on the effect of PGS on the probability of live birth after IVF.

Methods
PubMed and trial registers were searched for RCTs on PGS. Trials were assessed following predetermined quality criteria. The primary outcome was live birth rate per woman, secondary outcomes were ongoing pregnancy rate, miscarriage rate, multiple pregnancy rate and pregnancy outcome.

Results
Nine RCTs comparing IVF with and without PGS were included in our meta-analysis. Fluorescence in-situ hybridization was used in all trials and cleavage stage biopsy was used in all but one trial. PGS significantly lowered live birth rate after IVF for women of advanced maternal age (risk difference -0.08; 95% Confidence Interval -0.13 to -0.03). For a live birth rate of 26% after IVF without PGS, the rate would be between 13 and 23% percent using PGS. Trials where PGS was offered to women with a good prognosis and to women with repeated implantation failure suggested similar outcomes.

Conclusions
There is no evidence of a beneficial effect of PGS as currently applied on the live birth rate after IVF. On the contrary, for women of advanced maternal age PGS significantly lowers the live birth rate. Technical drawbacks and chromosomal mosaicism underlie this inefficacy of PGS. New approaches in the application of PGS should be evaluated carefully before their introduction into clinical practice.
In 1990, twelve years after the birth of the first child that was conceived by IVF (Steptoe and Edwards, 1978), the first pregnancies were announced that were obtained after transfer of embryos that had been tested for a specific genetic disorder (Handyside et al., 1990). In couples known to be at risk for a genetic disease, only unaffected embryos were transferred after IVF, aspiration of a single cell at the six- to eight-cell stage, and subsequent genetic analysis of the aspirated cell, in order to prevent affected children. This new technology was termed PGD (Sermon et al., 2004). At first, only DNA amplification techniques were used for this analysis, but not much later fluorescence in situ hybridization (FISH) techniques were introduced, allowing the detection of abnormal copy numbers of chromosomes, or aneuploidies (Griffin et al., 1992; Delhanty et al., 1993). Further technical developments made it possible to simultaneously analyse more than two chromosomes with the use of FISH (Munne et al., 1993).

In 1995 the first deliveries were reported after transfer of embryos that had been screened for aneuploidies (Verlinsky et al., 1995). The underlying rationale for this screening was no longer prevention of a genetic disease, but an expected increase in live birth rates after IVF, because embryos containing aneuploidies were thought not to implant or develop to term and hence to contribute to low live birth rates in specific groups of patients. This was called preimplantation genetic screening (PGS) (Wilton, 2002).

The beneficial effect of PGS was expected to be greatest in women of advanced maternal age, since aneuploidies in clinically recognized pregnancies occur more frequently when a woman passes 35 years of age (Hassold and Hunt, 2001) and it is in these women that pregnancy chances decline sharply both in normal conception and after IVF (Lintsen et al., 2007). Next to women of advanced maternal age, PGS has been offered to women with a history of recurrent miscarriage, women with a history of repeated implantation failure (i.e. several failed IVF cycles), and women with a partner with low sperm quality (severe male factor), mainly since high percentages of aneuploidies have been found in the embryos of these women (Munne et al., 1995; Marquez et al., 2000; Werlin et al., 2003; Silber et al., 2003; Munne et al., 2004; Kahraman et al., 2004; Wilding et al., 2004; Platteau et al., 2005; Rubio et al., 2005; Baart et al., 2006). More recently PGS has been offered to younger women (under 35 years of age) too, as high aneuploidy rates were found in their embryos as well (Baart et al., 2006; Goossens et al., 2009).

There has been an increase in the use of PGS in the last decade. The European Society for Human Reproduction and Embryology (ESHRE) PGD Consortium reported on 116 cycles of PGS performed worldwide in 1997–1998, steadily increasing to 3900 cycles in 2006 (Goossens et al., 2009). This number is probably an underestimation of the actual number of PGS cycles performed, as this report deals with a voluntary data collection. The only other data collection is a survey among all US-based infertility centers showing that 2197 cycles of PGS were performed in the USA alone in 2005 (Baruch et al., 2008; Goossens et al., 2008a). The efficacy of PGS is, however, still debated.

Here we present a systematic review and meta-analysis of RCTs dealing with the effect of PGS
on the probability of live birth after IVF.

**Methods**

**Search strategy**

We searched PubMed (www.pubmed.gov) and databases for registration of RCTs (www.clinicaltrials.gov) in December 2009 using the following search criteria: “(preimplantation genetic screening OR PGS OR preimplantation genetic diagnosis OR PGD OR aneuploidy screening OR PGD-AS OR screening for aneuploidies OR preimplantation testing OR embryo screening) AND (randomized trial OR randomised trial OR randomized study OR randomised study)”. The reference lists of included studies were searched for relevant studies.

**Study selection**

Two authors (SM, MT) independently examined the electronic search results for reports of possibly relevant trials and those reports were retrieved in full. Published RCTs were eligible for inclusion if they compared women undergoing IVF or ICSI with and without PGS. All trials were assessed following predetermined quality criteria. Validity was assessed in terms of method of randomization, presence of a power calculation, unit of analysis used, use of an intention-to-treat analysis and presence or absence of blinding. If necessary, we contacted the corresponding author of a report in an attempt to retrieve missing data. Trials were grouped based on the indications for PGS: advanced maternal age, recurrent miscarriage, repeated implantation failure, severe male factor, and good prognosis patients.

**Statistical analysis**

The effect of PGS was assessed for each indication separately. The primary outcome was live birth rate per woman, secondary outcome measures were ongoing pregnancy rate, miscarriage rate, multiple pregnancy rate and pregnancy outcome per woman. We calculated risk differences (RD) with 95% confidence intervals (CI) for each individual trial. The fixed-effect model was used to combine data for each indication separately and to combine all included studies. Statistical heterogeneity between results of studies was examined by inspecting the scatter in the data points on the graphs and the overlap of CIs, and by checking the I² statistic. A value of ≥ 50% was considered to indicate substantial heterogeneity. In case of substantial heterogeneity the random effects model was used instead of the fixed-effect model. Data were analyzed according to the intention-to-treat principle. Revman Software (Version 5, The Cochrane Collaboration) was used to combine data for meta-analysis.

**Results**

**Results of the search**

The literature search resulted in the identification of fourteen potentially eligible trials (Fig. I) (Gianaroli et al., 1997; Gianaroli et al., 1999; Werlin et al., 2003; Staessen et al., 2004; Stevens et al., 2004; Mastenbroek et al., 2007b; Staessen et al., 2008; Hardarson et al., 2008;
Mersereau et al., 2008; Blockeel et al., 2008; Jansen et al., 2008; Meyer et al., 2009; Debrock et al., 2009; Schoolcraft et al., 2009a).

Five trials were excluded after further assessment. Two trials were excluded since couples were allocated to the treatment or control group on the basis of their volunteer decision, instead of random allocation (Gianaroli et al., 1997; Gianaroli et al., 1999). PGS in these trials was performed for advanced maternal age and repeated IVF failure. One trial was excluded (Stevens et al., 2004) since there was data overlap with another publication (Schoolcraft et al., 2009a). One study was excluded since it contained insufficient data to assess the methodological quality (Werlin et al., 2003). In this trial PGS was performed for advanced maternal age, recurrent miscarriage and repeated implantation failure. Baseline characteristics of patients were not provided and method of randomization, concealment of allocation and blinding status were not mentioned. This trial reported no outcomes beyond preclinical pregnancy. One trial was excluded since it was an ongoing trial and since data were reported only as percentages without mentioning the unit of analysis (Mersereau et al., 2008). PGS was offered in this trial to women with a good prognosis.

Nine trials, representing 1589 treated women, were included in the meta-analysis (Staessen et al., 2004; Mastenbroek et al., 2007b; Staessen et al., 2008; Hardarson et al., 2008; Blockeel et al., 2008; Jansen et al., 2008; Meyer et al., 2009; Debrock et al., 2009; Schoolcraft et al., 2009a).

In addition, three RCTs that are currently recruiting patients were identified by searching www.clinicaltrials.gov. In two trials (NCT00646893; NCT00795795) PGS is offered to women of advanced maternal age. In one trial (NCT00547781) PGS is offered to women with repeated implantation failure. In these three trials biopsy will take place at cleavage stage (day 3 of embryo development) and FISH will be used for the analysis. One trial for advanced maternal age will apply assisted hatching to the embryos in the control group, the other two trials use IVF without PGS or assisted hatching as control treatment.
<table>
<thead>
<tr>
<th>Study</th>
<th>Patients</th>
<th>Cycles</th>
<th>Indication for PGS</th>
<th>Inclusion criteria</th>
<th>ART &amp; laboratory</th>
<th>Biopsy and FISH</th>
<th>Quality features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staessen 2004</td>
<td>389</td>
<td>389</td>
<td>AMA</td>
<td>≥ 37 years, normal karyotype of both partners, need for ICSI with motile sperm</td>
<td>ICSI; max 3 embryos for transfer if age 37-39 max 6 embryos if age ≥ 40</td>
<td>laser; mostly 2 blastomeres removed; biopsy on embryos with at least 5 blastomeres and with a maximum of 50% fragmentation; X, Y, 13, 16, 18, 21, 22</td>
<td>method of randomisation nr; concealment of allocation; not blind; power calculation performed; two centers; full paper; no intention to treat</td>
</tr>
<tr>
<td>Mastenbroek 2007</td>
<td>408</td>
<td>836</td>
<td>AMA</td>
<td>≥ 35-41 years, no previous failed IVF cycles, no objection to DET</td>
<td>IVF &amp; ICSI; max 2 embryos for transfer</td>
<td>laser; mostly 1 blastomere removed; biopsy on embryos with at least 4 blastomeres and with a maximum of 50% fragmentation; X, Y, 13, 16, 17, 18, 21</td>
<td>randomisation by computer; concealment of allocation; double blind; power calculation performed; four centers; full paper; analysis on intention to treat basis; data on cryo-cycles present</td>
</tr>
<tr>
<td>Hardarson 2008</td>
<td>109</td>
<td>109</td>
<td>AMA</td>
<td>≥ 38 years, at least three embryos of good morphological quality if DET, at least two embryos of good morphological quality if SET</td>
<td>IVF &amp; ICSI; max 2 embryos for transfer</td>
<td>laser; mostly 1 blastomere removed; biopsy on embryos with at least 6 blastomeres and with a maximum of 20% fragmentation; X, Y, 13, 16, 18, 21, 22</td>
<td>randomisation by computer; concealment of allocation nr; not blind; power calculation performed; two centers; full paper; intention to treat nr</td>
</tr>
<tr>
<td>Schoorkraft 2009</td>
<td>62</td>
<td>62</td>
<td>AMA</td>
<td>≥ 35 years, presence of at least 5 embryos with ≥ 6 cells and ≤ 15% fragmentation on day 3</td>
<td>IVF &amp; ICSI; max number of embryos transferred nr</td>
<td>laser or acidic tyrosides; mostly 1 blastomere removed; embryos for biopsy nr.; X, Y, 13, 15, 16, 17, 18, 21, 22</td>
<td>randomisation by computer; concealment of allocation nr; blindness nr; power calculation nr; centers nr; full paper; intention to treat nr</td>
</tr>
<tr>
<td>Debrock 2009</td>
<td>94</td>
<td>94</td>
<td>AMA</td>
<td>≥ 35 years with at least two fertilised oocytes available on day 1 after oocyte retrieval, and with at least two embryos consisting of six or more cells at day 3 after oocyte retrieval</td>
<td>IVF, ICSI &amp; TESE; max 2-3 embryos transferred, after July 2003 max 1 embryo transferred in 1st cycle of patients &lt;36 years</td>
<td>laser; 2 blastomeres removed; biopsy on embryos with at least six blastomeres; X, Y, 13, 16, 18, 21, 22</td>
<td>randomisation by blinded envelopes; concealment of allocation nr; not blind; power calculation nr; single center; abstract; no intention to treat data on cryo-cycles present</td>
</tr>
<tr>
<td>Staessen 2008</td>
<td>240</td>
<td>214</td>
<td>good prognosis</td>
<td>&lt; 36 years, need for ICSI with motile sperm, both partners normal karyotype</td>
<td>ICSI; single embryo transfer</td>
<td>laser; mostly 1 blastomere removed; biopsy on embryos with at least five blastomeres with a maximum of 50% fragmentation; X, Y, 13, 16, 18, 21, 22</td>
<td>randomisation by computer; concealment of allocation nr; not blind; power calculation performed but study was stopped after interim analysis because of futility; one center; full paper; intention to treat nr</td>
</tr>
<tr>
<td>Jensen 2008</td>
<td>101</td>
<td>101</td>
<td>good prognosis</td>
<td>&lt; 38 years, agreement to SET, first or second IVF, no cancelled cycles because of poor response, ≥ 8 follicles ≥ 1 cm on day 8-10 of stimulation, ≥ 4 embryos ≥ 7 cells on day 3, ≥ 2 blastoysts for biopsy</td>
<td>IVF/ICSI nr; single embryo transfer</td>
<td>laser; 2-9 trophectoderm cells removed; biopsy only on blastoysts; X, Y, 13, 18, 21</td>
<td>randomisation by sealed envelopes; concealment of allocation nr; blindness nr; power calculation performed but not reached; one center; full paper; intention to treat</td>
</tr>
<tr>
<td>Meyer 2009</td>
<td>47</td>
<td>47</td>
<td>good prognosis</td>
<td>&lt; 39 years, normal ovarian reserve, body mass index 30 kg/m², no smoking history, no hydrocephaly, presence of ejaculated sperm, normal uterus, ≤ 2 previous failed IVF cycles, ≥ 4 embryos containing at least 5 cells with &lt; 40% fragmentation</td>
<td>ICSI; max embryos for transfer nr</td>
<td>acidic tyrosides; mostly 1 blastomere removed; embryos for biopsy nr.; X, Y, 13, 16, 17, 18, 21, 22</td>
<td>randomisation by computer; concealment of allocation nr; blindness nr; power calculation performed but study was stopped after interim analysis because of large difference between the two groups; one center; full paper; no intention to treat</td>
</tr>
<tr>
<td>Blockeel 2008</td>
<td>162</td>
<td>139</td>
<td>RIF</td>
<td>≥ 35 years, normal karyotype of both partners, motile sperm</td>
<td>ICSI; max 3 embryos for transfer</td>
<td>method for zona drilling nr; between 2003 and Oct. 2005 two blastomeres were removed from embryos with at least six blastomeres, from Nov. 2005 onwards one blastomere was removed ,X, Y, 13, 16, 18, 21, 22</td>
<td>randomisation by computer; no concealment of allocation; blindness nr; power calculation performed; one center; full paper; intention to treat nr</td>
</tr>
</tbody>
</table>

AMA = advanced maternal age; DET = double embryo transfer; nr = not reported; RIF = repeated implantation failure; SET = single embryo transfer; TESE = testicular sperm extraction
Trial characteristics

Main characteristics and quality features of the nine included trials are presented in Table 1. In five studies PGS was offered to women of advanced maternal age (Staessen et al., 2004; Mastenbroek et al., 2007b; Hardarson et al., 2008; Debrock et al., 2009; Schoolcraft et al., 2009a). In three studies PGS was offered to good prognosis patients (Staessen et al., 2008; Jansen et al., 2008; Meyer et al., 2009). In one study PGS was offered to women with repeated implantation failure (Blockeel et al., 2008).

In one trial a maximum of three treatment cycles was offered to each participating couple (Mastenbroek et al., 2007b). In one trial a couple could participate several times with independent randomization for each cycle (Debrock et al., 2009). All other trials offered one treatment cycle to each participating couple. One trial analyzed two to nine trophectoderm cells from the blastocyst stage of embryo development, aspirated from the embryo at the fifth or sixth day after fertilization (Jansen et al., 2008). All other trials analyzed blastomeres from the cleavage stage of embryo development, aspirated from the embryo at the third day after insemination. In the majority of these trials preferably one blastomere was aspirated for analysis, and in two trials two blastomeres were aspirated (Staessen et al., 2004; Debrock et al., 2009). All trials used FISH for the analysis of the aspirated cells. There were no trials that analyzed all chromosomes. One embryo was transferred to the uterus in two trials (Staessen et al., 2008; Jansen et al., 2008), two embryos were transferred if these were available in two trials (Mastenbroek et al., 2007b; Hardarson et al., 2008), and up to three embryos were transferred in one trial (Blockeel et al., 2008). In one trial the number of embryos transferred depended on the age of the woman, since up to three embryos were transferred when the recipient was between 37 and 39 years old and up to six embryos were transferred when the recipient was 40 years or older (Staessen et al., 2004). In one trial a maximum of three embryos were transferred in the first period of the trial and only one embryo in the second part of the trial for first treatment cycles of women younger than 36 years (Debrock et al., 2009). In two trials transfer policy was not described (Meyer et al., 2009; Schoolcraft et al., 2009a). In one trial there was a difference in embryo transfer policy between the PGS treatment group and the control group. In the PGS treatment group embryo transfer was performed on day five while in the control group transfer was performed on day three (Hardarson et al., 2008). Data on cryopreservation cycles and pregnancies resulting from those cycles were available from two trials (Mastenbroek et al., 2007b; Debrock et al., 2009).

The included trials used embryo implantation rate (Staessen et al., 2004; Blockeel et al., 2008; Debrock et al., 2009), (clinical) pregnancy rate (Hardarson et al., 2008) ongoing pregnancy rate (Mastenbroek et al., 2007b) or live birth rate (Staessen et al., 2008; Jansen et al., 2008; Meyer et al., 2009) as primary outcome measure. For one trial the primary outcome was not specified (Schoolcraft et al., 2009a). Live birth rate, the primary outcome measure of our meta-analysis, was reported in nine trials (Table II) (Staessen et al., 2004; Mastenbroek et al., 2007b; Staessen et al., 2008; Hardarson et al., 2008; Blockeel et al., 2008; Jansen et al., 2008; Meyer et al., 2009; Debrock et al., 2009; Schoolcraft et al., 2009a). For two trials the authors kindly provided missing or additional data on live birth rates upon request (Staessen et al., 2004; Debrock et al., 2009). The ongoing pregnancy rate was available for five trials.
(Staessen et al., 2004; Mastenbroek et al., 2007b; Staessen et al., 2008; Blockeel et al., 2008; Debrock et al., 2009), the miscarriage rate was available for nine trials (Staessen et al., 2004; Mastenbroek et al., 2007b; Staessen et al., 2008; Hardarson et al., 2008; Blockeel et al., 2008; Jansen et al., 2008; Meyer et al., 2009; Debrock et al., 2009; Schoolcraft et al., 2009a), multiple pregnancy rate was available in six trials (Staessen et al., 2004; Mastenbroek et al., 2007a; Staessen et al., 2008; Blockeel et al., 2008; Debrock et al., 2008; Hardarson et al., 2008; Hardarson et al., 2008 personal communication), and information on congenital malformations was available from three trials (Mastenbroek et al, 2007; Staessen et al., 2004 personal communication; Hardarson et al., 2008 personal communication).

Power calculation was used to calculate sample size in seven trials (Staessen et al., 2004; Mastenbroek et al., 2007b; Staessen et al., 2008; Hardarson et al., 2008; Blockeel et al., 2008; Jansen et al., 2008; Meyer et al., 2009). Two trials were terminated prematurely since continuing was considered not useful after interim analysis (Staessen et al., 2008; Hardarson et al., 2008). One trial was terminated prematurely since interim analysis showed such a large difference that it was considered unethical to continue (Meyer et al., 2009), and one trial was terminated prematurely since the trend was opposite to what was expected and cryopreserved blastocysts were accumulating in the control group (Jansen et al., 2008). In two studies the power calculation was based on embryos instead of women (Staessen et al., 2004; Blockeel et al., 2008). This results in a underpowered sample size when analyzing an outcome where women are the unit of analysis, as is done in this meta-analysis by using live birth rate per woman as primary outcome (Mastenbroek et al., 2005).

**Extracted data**

In one trial randomization assigned 200 women to the PGS treatment group and 200 women to the control group (Staessen et al., 2004). However, in the PGS group one woman did not fulfill the inclusion criteria, and in the control group 10 women did not, so therefore we included 199 and 190 women, respectively, in our meta-analysis (Table II). In one trial, randomization assigned 95 women to the PGS treatment group and 105 women to the control group (Blockeel et al., 2008): among these were eight and ten women, respectively, that were not fulfilling the inclusion criteria and we excluded those from the meta-analysis. The authors excluded 15 women in the PGS group and 28 women in the control group owing to ‘insufficient ovarian response’, ‘stop further fertility treatment’, ‘wrong allocation’ and ‘spontaneous pregnancy’. No results were provided for these women, so we could not include them in our meta-analysis following the intention to treat principle. Thus we included 72 and 67 women for the PGS and control group, respectively, for this trial in our meta-analysis. One trial reported only percentages and used transfers and pregnancies as units of analysis, therefore we recalculated the numbers per patient for the various outcomes (Schoolcraft et al., 2009a). One trial included drop-outs in their analysis (Mastenbroek et al., 2007b), four trials did not or did not correctly do so (Staessen et al., 2004; Staessen et al., 2008; Meyer et al., 2009; Debrock et al., 2009) and for the other trials nothing was mentioned about drop-outs. Missing information about drop-outs was kindly provided on request by the authors of one trial (Debrock et al., 2009). We analyzed all trials following the intention to treat principle. All drop-outs, where available, were included in our analysis.
For the primary outcome live birth rate we combined the outcomes of five trials, reporting on 537 women in the PGS group and 525 women in the control group (Fig. II). Live birth rate was significantly lower after PGS (18%) compared to the control group (26%) (N= 1062; RD -0.08; 95% CI -0.13 to -0.03). This suggests that for this group of women with a live birth rate of 26% after IVF without PGS, the rate would be between 13 and 23% when using PGS. There was no indication for statistical heterogeneity.

Ongoing pregnancy rate was significantly lower in the PGS group (18%) than in the control group (26%) (N=891; RD -0.08, 95% CI -0.13 to -0.02) without indication for statistical heterogeneity. The miscarriage rate was not significantly different between the PGS group (11%) and the control group (12%) (N=1062; RD -0.01, 95% CI -0.05 to 0.02) without indication for statistical heterogeneity.

The multiple pregnancy rate was not significantly different between the PGS and the control group; 14 twin and one triplet pregnancies were reported in the PGS group, with one twin pregnancy resulting from a vanishing triplet pregnancy, and 23 twin and three triplet pregnancies were reported in the control group. Two triplet pregnancies were reduced to a singleton pregnancy and one to a twin pregnancy, which ended in a singleton pregnancy. This results in an RD of 0.00 (N=199; 95% CI -0.12 to 0.11; no indication for statistical heterogeneity).

In the PGS group one trisomy 18 (after spontaneous conception) was reported, one intrauterine death (due to abruptio placentae) and one premature delivery of twins at 24 weeks of gestation, resulting in the postpartum death of both children (Mastenbroek et al., 2007b). In the control group one trisomy 21 was reported (Staessen et al., 2004, personal communication), one trisomy 18, one conception with a cleft lip and palate and one intrauterine death of a fetus (Mastenbroek et al., 2007b). In one trial no congenital malformations were detected (Hardarson et al., 2008 personal communication). The other studies did not report on pregnancy outcome.
For the primary outcome live birth rate we combined the outcomes of three trials, reporting on 198 women in the PGS group and 190 women in the control group (Fig. II). The live birth rate was lower after PGS (32%) compared to the control group (42%) (N=388; RD -0.10, 95% CI -0.20 to -0.01), however there was substantial statistical heterogeneity (I²= 76%). When pooling the data using the random effects model because of this substantial statistical heterogeneity, no difference in live birth rate could be proven (N=388; RD -0.17, 95% CI -0.39 to 0.04). This suggests that for this group of women with a live birth rate of 42% after IVF without PGS, the rate would be between 3 and 46% when using PGS.

Ongoing pregnancy rate was only reported in one trial (Staesssen et al., 2008), therefore no meta-analysis could be performed. The ongoing pregnancy rate was not significantly different between the PGS (31%) and the control group (33%) (N=240; RD -0.02, 95% CI -0.13 to 0.10). The combined miscarriage rate was 10% in the PGS group and 9% in the control group (N=388; RD 0.01, 95% CI -0.04 to 0.07). There was substantial statistical heterogeneity (I²= 69%). Pooling the data using the random effect model showed no difference in miscarriage.

### Good prognosis patients

For the primary outcome live birth rate we combined the outcomes of three trials, reporting on 198 women in the PGS group and 190 women in the control group (Fig. II). The live birth rate was lower after PGS (32%) compared to the control group (42%) (N=388; RD -0.10, 95% CI -0.20 to -0.01), however there was substantial statistical heterogeneity (I²= 76%). When pooling the data using the random effects model because of this substantial statistical heterogeneity, no difference in live birth rate could be proven (N=388; RD -0.17, 95% CI -0.39 to 0.04). This suggests that for this group of women with a live birth rate of 42% after IVF without PGS, the rate would be between 3 and 46% when using PGS.
rate (N=388; RD 0.05, 95% CI -0.06 to 0.16).

In one trial three multiple pregnancies occurred even though single embryo transfer was applied (Staessen et al., 2008); one in the PGS group and two in the control group (RD -0.03; 95% CI -0.12 to 0.06). In one trial multiple pregnancy rate was not mentioned (Jansen et al., 2008) and in one trial it was only mentioned as a percentage and it was not possible to calculate this into exact numbers (Meyer et al., 2009).

In one trial no major or minor congenital abnormalities were detected (Meyer et al., 2009). The other two trials did not report on congenital abnormalities.

Repeated implantation failure
The single trial in which PGS was offered to women with repeated implantation failure, showed a significant difference in live birth rate between the PGS group (21%) and the control group (39%) (N=139; RD -0.18, 95% CI -0.33 to -0.03) (Fig. II). This suggests that for this group of women with a live birth rate of 39% after IVF without PGS, the rate would be between 6 and 36% when using PGS.

The ongoing pregnancy rate per woman was significantly different between the PGS group (21%) and the control group (39%) (N=139; RD -0.18, 95% CI -0.33 to -0.03). The miscarriage rate was not significantly different between the PGS (14%) and the control group (7%) (RD 0.06, 95% CI -0.04 to 0.17). In the PGS group three twin pregnancies were reported and in the control group 10 twin pregnancies were reported (RD -0.18; 95% CI -0.46 to 0.09). All children in the PGS group as well as the control group were born healthy.

Developmental stage
There was only one trial that analysed embryos at blastocyst stage (Jansen et al. 2008). This was a trial with good prognosis patients and live birth rates were significantly lower after PGS (N= 101; RD -0.22, 95% CI -0.41 to -0.03). When PGS was performed at the cleavage stage for the same indication the live birth rate was also lower after PGS, but this difference was not significant (N= 287; Fixed-effect model: RD -0.06, 95% CI -0.17 to 0.05, I2 84%; Random effects model: RD -0.16, 95% CI -0.52 to 0.19).

Method of analysis
All trials used FISH for the analysis of aspirated cells, thus no comparison could be made with other methods of analysis.

Discussion

Interpretation
Our review and analysis of the literature clearly shows that there is no evidence of a beneficial effect of PGS as currently conducted on live birth rates after IVF. In women of advanced maternal age PGS significantly lowered live birth rate after IVF (RD -0.08; 95% CI -0.13 to
In women with a good prognosis (RD -0.17, 95% CI -0.39 to 0.04) and in women with repeated implantation failure (RD -0.18, 95% CI -0.33 to -0.03) a similar effect after PGS was seen.

It has been argued that the inclusion criteria for the trials included in this review were not targeted specifically enough to women with a high risk of embryonic aneuploidy (Handyside and Thornhill, 2007). Data from one of the trials included in this review was analysed to assess whether the effect of PGS indeed differs with variable risk for embryonic aneuploidy, but no beneficial effect of PGS was found regardless of the risk for embryonic aneuploidy (Twisk et al., 2008). It has also been argued that the trials included in this review lack technical prowess causing them to be neither valid nor generalizable (Cohen and Munne, 2005; Cohen and Grifo, 2007; Munne et al., 2007; Simpson, 2008; Munne et al., 2010), but all trials from multiple independent established groups showed the same negative effect of PGS and therefore it seems particularly justified to generalize the outcome of these studies and to conclude that there is no beneficial effect of PGS in terms of increased live birth rates.

Reasons for inefficacy

There are several possible explanations for the inefficacy of PGS. These concern technical aspects of the PGS procedure, such as possible harm from the biopsy procedure, the failure rate of the technique, and the limitations of the FISH analysis, in addition to an intrinsic biological feature of the analysed embryos, which is their mosaic nature.

Technical aspects

**Biopsy**

A biopsy entails taking an embryo out of the incubator for at least a couple of minutes, making a hole in the zona pellucida -by mechanical dissection, by using acidic Tyrode’s solution or by using a laser- and aspirating one or more cells (De Vos and Van Steirteghem, 2001). It seems plausible to assume that an embryo would at least not benefit from this, but the extent to which this could be harmful for an embryo is still a moot point as data on this topic are limited.

It has been shown that aspiration of cells at the two to four cell stage, or the removal of a quarter of the cell mass, influences embryo development by lowering the ratio of inner cell mass cells to trophectoderm cells (Tarin et al., 1992). Embryo development was reported to be less affected when biopsy was performed at the eight cell stage (Hardy et al., 1990; Hardy and Handyside, 1993). However, the effect of biopsy alone on pregnancy rates has never been properly studied.

It has been argued that the removal of two blastomeres at day three of development is likely to negatively influence pregnancy rates (Cohen et al., 2007), similar to the negative effect of cell loss after cryopreservation on pregnancy rates (Edgar et al., 2000a; Edgar et al., 2000b), which would be relevant to two trials included in our review (Staessen et al., 2004; Debrock et al., 2009). A prospective controlled trial showed aspiration of two blastomeres instead of one to negatively influence embryo development without an effect on live birth rates (Goossens et al., 2008b), although more recently a cohort study by the same research group did show
a detrimental effect of the aspiration of two blastomeres on live birth rates (De Vos et al., 2009).

**Failure rate**

PGS has not been without failure. The success rates for the separate steps of PGS, i.e. biopsy, fixation and FISH-analysis, seem to be similar between the largest RCTs included in this review: 97.2 and 94.2% successfully biopsied embryos, 89.1 and 93.6% successfully fixated blastomeres, and 92.3 and 92.9% blastomeres with successful FISH analysis for the centers in Amsterdam and Brussels, respectively (Michiels et al., 2006; Mastenbroek et al., 2007b): this led to 20.1% of embryos without a diagnosis in the former trial and 2.4% in the latter trial (instead of a cumulative 19.3%), since in the latter trial a second cell was available for the analysis of each embryo (Staessen et al., 2004; Mastenbroek et al., 2007b).

These trials have been fiercely criticized for their failure rates (Cohen and Munne, 2005; Cohen and Grifo, 2007; Munne et al., 2007). The ESHRE PGD consortium reports a failure rate of 8.8% per embryo for all indications of PGS and 9.9% for the indication advanced maternal age in their latest data collection (data collection I-IX (Goossens et al., 2009)). This could reflect underreporting, since the data collection is voluntary, not based on an intention to treat analysis, and incomplete reporting has previously been noted (Sermon et al., 2007). Another reason for the reported difference in diagnostic efficiency could be the way PGS was performed. Failure rate in the criticized trials was especially high in embryos with four or five blastomeres with up to 50% fragmentation (Mastenbroek et al., 2007a), possibly related to the arrested nature of these embryos and the resulting compromised DNA integrity. In more recent years it became common to perform biopsy only on better quality embryos, which obviously limits the applicability of PGS, but possibly increases diagnostic efficiency (Munne et al., 2003; Summers and Foland, 2009). One trial on advanced maternal age that biopsied only better quality embryos, i.e. embryos with at least six cells and less than 20% fragmentation, had a failure rate of 10.6% (Hardarson et al., 2008). Another trial on advanced maternal age that used even stricter inclusion criteria, i.e. the presence of at least five embryos with at least six or more cells and less than 15% fragmentation, had a failure rate of 5.0% (Schoolcraft et al., 2009a). The last trial included in our review for advanced maternal age performed two cell biopsy and diagnostic efficiency could not be corrected for this (Debrock et al., 2009).

Performing PGS on lower quality embryos thus seems to have lowered diagnostic efficiency in the two largest available trials; it is important to note that this does not necessarily translate to decreased PGS efficacy, as these lower quality embryos are simply excluded in the other trials. Indeed, removing the trials with the highest failure rates from our analysis, did not result in a different RD for our main outcome (RD -0.09; 95% CI -0.18 to 0.01 versus RD -0.08; 95% CI -0.13 to -0.03). This result is not surprising because statistical heterogeneity was zero % in our initial analysis. Other adjustments to the PGS technique that have been suggested in recent years, such as the reanalysis of a nucleus with a different probe set in case of a failed FISH analysis (Colls et al., 2007) or the use of an increased number of probes (Colls et al., 2009), could also contribute to lower PGS failure rates, but prospective comparative data to substantiate this are not available.
**FISH**

All trials included in our review used FISH for the analysis of aspirated cells. The intrinsic limitations of FISH could well in part underlie the inefficacy of PGS. Owing to technical limitations, i.e. the available number of fluorochromes and the time available for the analysis, one can only test for a limited number of chromosomes in PGS using FISH. PGS therefore fails to detect embryos that are aneuploid for the chromosomes that are not tested. In addition, it is not exactly known which combination of probes is best. The selection of probes is in fact quite arbitrary since there are no large detailed studies on the frequency of numerical abnormalities of all 23 chromosome pairs in cleavage stage embryos. Prospective studies comparing different probe sets in relation to live birth rates after PGS have not been conducted. Furthermore, scoring errors of FISH signals may arise from loss or damage of nuclear material, overlapping signals, split signals, diffused signals, hybridization failure and probe inefficiency. FISH analysis has a 92–99% accuracy per probe, so when using a multi-probe panel on one cell, the risk of misdiagnosis is significant (Ruangvutilert et al., 2000; Michiels et al., 2006; Deugarte et al., 2008). To illustrate this: an accuracy of 98% per probe will result in an estimated 15% error rate for an eight-probe panel \((0.98^8)\). The low positive predictive value of the test will result in the exclusion of embryos for consideration for transfer, thereby worsening PGS efficacy (Michiels et al., 2006; Deugarte et al., 2008; Treff et al., 2010; Northrop et al., 2010).

**Mosaicism**

Embryo mosaicism, the condition that the chromosomal constitution differs between blastomeres of the same embryo, could well be an important reason for the inefficacy of PGS. Clinical studies on the effect of mosaicism on PGS commonly focus on ‘misdiagnosis’ rates, i.e. false-negative and false-positive-rates of FISH analysis, rather than on the impact of mosaicism on pregnancy or live birth rates (Los et al., 2004; Staessen et al., 2004). It is diploid-aneuploid mosaicism, i.e. the condition that an embryo consists of both normal and aneuploid blastomeres, that could be causal to the lower live birth rates after PGS. Aspiration of aneuploid blastomeres from diploid-aneuploid embryos leads to the discarding of these embryos, although they are potentially viable as these embryos contain normal blastomeres. Conversely, aspiration of normal blastomeres from diploid-aneuploid embryos results in transfer or cryopreservation of these embryos, although the proportion of normal blastomeres was reduced, thereby potentially hampering the embryos potential. A recent systematic review of the literature reported that 46% of 107 developing cleavage stage embryos that were analyzed for eight or more chromosomes were diploid-aneuploid mosaic, with a mean of 47% normal blastomeres in these embryos (van Echten et al., submitted). This means that in 24% of all embryos \((0.46^*0.53)\) an aneuploid blastomere will be aspirated, thus leading to discarding of these potentially viable embryos and in 22% of all embryos \((0.46^*0.47)\) a normal blastomere is aspirated during the PGS procedure, thus leading to transfer or cryopreservation of embryos with a relatively lower number of normal blastomeres. It must be noted that these data were obtained from spare, and thus lower quality embryos after IVF, which have been suggested to contain higher percentages of abnormalities (Ziebe et al., 2003; Magli et al., 2007). Furthermore, it is unknown to what extent FISH errors contribute to the prevalence of mosaicism and thus to the percentages mentioned here (Treff et al., 2010), although recently
support for the above reasoning was obtained with the use of genome-wide comparative genome hybridization arrays (CGH-arrays) (Vanneste et al., 2009a; Vanneste et al., 2009b). Notwithstanding the uncertainty of the exact percentages, there seems to be a clear role for mosaicism in the inefficacy of PGS.

New developments
The publication of an increasing number of trials that were unable to show any benefit of PGS led to a renewed and increasing interest in further development of, and adjustments to, the PGS technique in recent years. The focus of these developments in technique optimization is on the timing of biopsy, i.e. the developmental stage of the embryo when PGS is performed, and on the method that is used to analyze the aspirated cells.

Developmental stage
Of the ten trials included in our review, nine performed PGS at cleavage stage (generally at day three of embryo development) and one performed PGS at the blastocyst stage (generally at day 5 or 6 of embryo development). Blastomere aspiration at day three has been most widely used in PGS (Goossens et al., 2009), because it is less harmful for the embryo than at earlier stages (Hardy et al., 1990; Tarin et al., 1992), and it provides enough time for the analysis of cells before transfer (De Vos and Van Steirteghem, 2001). Although the single trial included in our review that performed PGS at the blastocyst stage showed lower live birth rates after IVF/ICSI with PGS (Jansen et al., 2008), there is an increased interest in blastocyst biopsy. In blastocyst biopsy more cells can be aspirated and mosaicism seems to be less frequent (Bielanska et al., 2002). It is possible to remove only the extra-embryonic trophectoderm cells, leaving the inner cell mass, which contains the cells the future child derives from, intact (Dokras et al., 1990). However, it is unknown whether trophectoderm cells are representative of the cells the future child derives from and blastocyst PGS leaves only limited time for the analysis of the aspirated cells.

Although the first reported children after PGS were conceived with the use of polar body aspiration (Verlinsky et al., 1995; Verlinsky et al., 2004), its use remained limited compared to cleavage stage aspiration (Goossens et al., 2009). With the current evidence for inefficacy of PGS using cleavage stage aspiration, there is a renewed interest in the analysis of polar bodies for PGS purposes (Geraedts et al., 2009). An advantage is that both polar bodies are extra-embryonic material and are expected to have no biological role in the further development of the embryo. A disadvantage is that only maternal meiotic errors and no abnormalities from paternal origin or abnormalities arising during mitosis (mosaicism) can be evaluated (De Vos and Van Steirteghem, 2001; Kuliev and Verlinsky, 2004; Montag et al., 2009). Yet it is assumed that the extent of paternal contribution to embryonic aneuploidy is limited and surely less than the maternal contribution (Nicolaides and Petersen, 1998; Magli et al., 2009; Montag et al., 2009). A preclinical study recently conducted by the ESHRE PGS Task Force has shown that PGS using polar body biopsy and array-technology for the analysis of the polar bodies is technically feasible and now requires a prospective RCT to evaluate its efficacy (Geraedts, 2010).
Method of analysis

Optimising the FISH technique used for PGS has been proposed, such as the re-analysis of a cell with different FISH probes for the same chromosomes in case of FISH failure (Colls et al., 2007), or increasing the number of FISH probes by performing multiple rounds of FISH (Colls et al., 2009). Whether these modifications indeed improve PGS efficacy, in terms of live birth rate per women, has not yet been shown. Alternatives to the use of FISH, such as methods based on CGH or the analysis of single nucleotide polymorphisms (SNPs), have also been proposed (Wells et al., 2008). The use of CGH-based methods for PGS is not new; the first report of a child born after IVF with PGS using CGH stems from 2001 (Wilton et al., 2001). As the time needed for the CGH analysis (three to four days) exceeds the time available between biopsy and transfer (one to two days), cryopreservation of all embryos was necessary, which probably limited its widespread use because cryopreservation methods hampered the embryos potential for implantation, already compromised by the biopsy (Edgar et al., 2000b; Wilton et al., 2001; Stachecki et al., 2005; Magli et al., 2006). With recent advances in cryopreservation technology, the need for cryopreservation of embryos is considered to be less of a problem, resulting in an increased interest in PGS with the use of CGH (Kuwayama, 2007; Teramoto and Kato, 2007; El-Toukhy et al., 2009; Keskinetepe et al., 2009; Schoolcraft et al., 2009b). At the same time, increased effort is put in the development of array-based methods that allow for analysis within the available time-frame between biopsy and transfer (Hellani et al., 2008). Cryopreservation is thus not required, and these methods potentially have the advantage of being less complex and easier to use than the above mentioned CGH analysis. Some of these array-based methods use the principle of CGH, but now labeled DNA is hybridized with DNA spotted on a slide instead of hybridization with metaphase chromosomes, with the potential benefit of automated analysis and less time needed for the hybridization (Wells et al., 2008). Other array-based methods analyze SNPs, i.e. single nucleotide variations in the human genome to detect aneuploidies (Wells et al., 2008; Treff et al., 2010; Northrop et al., 2010).

Currently these developments are at the stage of method assessment or at the stage of pilot studies. No trials that investigate the value of these new approaches in PGS have yet been registered. Nonetheless, these developments have already been implemented in clinical practice at least in some clinics. Indeed, the theory behind these developments sounds plausible, the techniques used are attractive and provide potential, and the first results seem promising. However, a decade ago, this was exactly the same when PGS was developed and introduced into clinical practice. Any fair-minded person will conclude from this review that PGS as widely practiced has for many women been a significant additional cost without robust evidence of any benefit, and for some their chance of a live birth has probably been harmed. The same issues that are suggested to undermine PGS as reviewed here, i.e. possible harm of the biopsy procedure, a high failure rate, incorrect diagnosis by the technique used for analysis, and mosaicism, could very well also undermine these new approaches. Therefore, new approaches to PGS should not simply be introduced into clinical practice; it is essential that method assessment and pilot studies must first provide sufficient evidence to allow rigorous RCTs on the efficacy of these new techniques.
Conclusions
There is no evidence of a beneficial effect of PGS as currently applied on the live birth rate after IVF. On the contrary, for women of advanced maternal age PGS significantly lowers the live birth rate. New approaches in the application of PGS should be evaluated carefully before their introduction into clinical practice.

References


Cohen, J., Wells, D., and Munne, S. (2007) Removal of 2 cells from cleavage stage embryos is likely to reduce the efficacy of chromosomal tests that are used to enhance implantation rates. Fertil. Steril. 87, 496-503.


embryonic development and clinical outcome after the biopsy of one or two blastomeres for preimplantation genetic diagnosis. Hum. Reprod. 23, 481-492.


A SYSTEMATIC REVIEW AND META-ANALYSIS OF RCTs


