Preimplantation genetic screening: a reappraisal
Mastenbroek, S.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 8

Summary
In the majority of in vitro fertilization (IVF) cycles multiple embryos are created after ovarian hyperstimulation. The viability of these embryos, and, as a consequence, the chances for an embryo to successfully implant, is subject to biological variation. To achieve the best possible live birth rates after IVF while minimizing the risk for multiple pregnancy, one or two embryos that are considered to have the best chances of implanting are selected for transfer. Subsequently, supernumerary embryos with a good chance of implanting are selected for cryopreservation and possible transfer in the future while remaining embryos of poor quality are discarded.

Since the earliest days of IVF, the method of choice for embryo selection has been morphological evaluation, but with this type of embryo selection implantation rates in general do not exceed 35 percent (Ebner et al., 2003; Gerris, 2005; Centers for Disease Control and Prevention et al., 2010). This has resulted in a strong drive for finding alternative selection methods.

Preimplantation genetic screening (PGS) has been proposed as one such alternative method to select embryos for transfer in an IVF treatment. In PGS, a single blastomere is aspirated from each embryo, and the copy number of a set of chromosomes is then determined in that blastomere. Embryos that are identified as abnormal are then discarded, and embryos with a normal genetic constitution are selected for transfer.

The underlying rationale for this screening was an expected increase in live birth rates after IVF, because numerical chromosomal abnormalities were known to exist in human preimplantation embryos (Angell et al., 1983), and transferred embryos containing these abnormalities were thought not to implant or develop to term and hence to contribute to low live birth rates in specific groups of patients (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 also been offered to younger women (under 35 years of age), as high aneuploidy rates were also found in their embryos (Baart et al., 2006; Goossens et al., 2009).

In 1995, the first deliveries were reported after transfer of embryos that had been screened for aneuploidies (Verlinsky et al., 1995). At the time of starting the studies described in this thesis, the use of PGS had become increasingly common, in particular among women of advanced maternal age (Verlinsky et al., 2004; Sermon et al., 2007). It was even suggested
that PGS would become a standard procedure for all women undergoing IVF (Verlinsky et al., 2004), but evidence supporting the use of PGS was limited.

Observational studies comparing IVF with and without PGS had shown that PGS was associated with higher implantation rates per transferred embryo but not with an increase in the rate of ongoing pregnancies per initiated cycle or per follicular aspiration (Gianaroli et al., 1999; Munne et al., 1999; Obasaju et al., 2001; Munne et al., 2003; Montag et al., 2004).

In view of this lack of robust evidence, we conducted a multicenter, double-blind, randomized controlled trial comparing three cycles of IVF with and without PGS in women 35 through 41 years of age (Chapter 2). The primary outcome measure was ongoing pregnancy at 12 weeks of gestation. Four hundred eight women (206 assigned to PGS and 202 assigned to the control group) underwent 836 cycles of IVF (434 cycles with and 402 cycles without PGS).

The ongoing-pregnancy rate was significantly lower in the women assigned to PGS (52 of 206 women (25%)) than in those not assigned to PGS (74 of 202 women (37%); rate ratio, 0.69; 95% confidence interval (CI),0.51 to 0.93). The women assigned to PGS also had a significantly lower live-birth rate (49 of 206 women (24%) versus 71 of 202 women (35%); rate ratio, 0.68; 95% CI, 0.50 to 0.92).

Thus, PGS did not increase but instead significantly reduced the rates of ongoing pregnancies and live births after IVF in women of advanced maternal age. These results argued strongly against routinely performing PGS as an adjunct to IVF in this group of women.

There are several possible explanations for this inefficacy of PGS. These concern technical aspects of the PGS procedure, such as possible harm to the embryo from the biopsy procedure, the failure rate of the technique, the limitations of the FISH analysis, and the intrinsic biological feature of chromosomal mosaicism in the analyzed embryos.

Mosaicism in human preimplantation embryos is the condition that the chromosomal constitution differs between blastomeres of the same embryo (Delhanty et al., 1993). Especially diploid-aneuploid mosaicism, i.e. the condition that an embryo consists of both diploid and aneuploid blastomeres, 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 is reduced, thereby potentially hampering the potential of the embryos to implant.

We analyzed 360 embryos from our randomized controlled trial on PGS, using fluorescence in situ hybridisation (FISH) analysis for chromosomes 1, 13, 16, 17, 18, 21, X and Y, to investigate whether mosaicism indeed hampers the efficacy of PGS (Chapter 3). Fifty-five percent of all analyzed embryos contained both diploid and aneuploid blastomeres and 46 percent of

SUMMARY
the blastomeres of these diploid-aneuploid mosaic embryos were diploid. We then used the chromosomal makeup of these embryos to construct a model to estimate the effect of chromosomal mosaicism on pregnancy rates after PGS, as clinical studies on the effect of mosaicism on PGS are scarce and 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 rates (Los et al., 2004; Staessen et al., 2004).

Our model showed that diploid-aneuploid mosaicism indeed provides a plausible explanation for the reduced pregnancy chances after PGS. We concluded that the high prevalence of diploid-aneuploid mosaicism precludes the determination of the chromosomal makeup of a human embryo by analyzing a single cell using FISH during cleavage stage development. To validate the mosaicism rates we found in our study, we subsequently performed a systematic review of the literature on mosaicism in human preimplantation embryos (Chapter 4). Similar to reviews based on individual patient data (Broeze et al., 2010), we only included studies that provided full information on each individual embryo. This allowed us to use one set of predetermined criteria to determine mosaicism rates. A total of 815 embryos from 36 studies could be classified according to our predetermined criteria. Of these, 177 (22%) were diploid, 599 (73%) were mosaic, of which 480 (59% of the total number of embryos) were diploid-aneuploid mosaic and 119 (14% of the total number of embryos) were aneuploid mosaic, and 39 (5%) contained other numerical chromosomal abnormalities. The distribution of the embryos over these categories was associated with the developmental stage of the embryos, the method used for analysis and the number of chromosomes analysed.

The data from this review confirmed the mosaicism rates found in the cohort of embryos that we had analyzed, and corroborated the conclusion that mosaicism is undermining the reliable determination of the ploidy status of a preimplantation embryo by analyzing a single cell.

Simultaneously, and especially after our randomized controlled trial on the efficacy of PGS, multiple other trials were conducted (Staessen et al., 2004; Blockeel et al., 2008; Hardarson et al., 2008; Jansen et al., 2008; Staessen et al., 2008; Meyer et al., 2009; Schoolcraft et al., 2009a; Debrock et al., 2010). Even after the publication of these trials, controversy on the effectiveness of PGS remained and a call for more randomized controlled trials was made by many (Cohen and Grifo, 2007; Munne et al., 2007; Harper et al., 2008; Jansen et al., 2008; Simpson, 2008; Yakin and Urman, 2008).

To put this into perspective, we wrote an opinion paper on the significance of performing even more randomized controlled trials on the efficacy of PGS (Chapter 5). We stated that, it would be unethical to perform additional randomized controlled trials for the indication advanced maternal age using cleavage stage biopsy and FISH in view of the lack of evidence for the effectiveness of PGS and the accumulating evidence for its harm.

We first performed a meta-analysis of published randomized comparative data on PGS for the indication advanced maternal age. We then calculated the necessary power of a new randomized controlled trial to shift the common odds ratio at that time from 0.56 to a
common odds ratio significantly above 1.0, indicative of a positive effect of PGS on ongoing pregnancy rates. Assuming a clinically relevant increase in ongoing pregnancy rate after PGS in this new RCT of 20%, this trial should then include at least 6000 cycles.

We then argued why it would be unethical to perform a trial of such magnitude. The first reason is that the most important ethical condition to perform any trial, i.e. the concept of being in equipoise, was no longer met (Lilford and Jackson, 1995; Lilford, 2003). The second reason is that there were intrinsic limitations of the then applied PGS techniques that would make it highly unlikely that a new trial would ever find an increase in ongoing pregnancies.

We concluded that, before embarking on new randomized controlled trials for indications other than advanced maternal age using cleavage stage biopsy and FISH, or new randomized controlled trials for any indication using emerging techniques, such as the analysis of all chromosomes using DNA-arrays, it is essential that pilot-studies 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 ten 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.

We then performed a systematic review and meta-analysis of randomized controlled trials on the effect of PGS on the probability of live birth after IVF for all indications of PGS (Chapter 6). The primary outcome was live birth rate; secondary outcomes were ongoing pregnancy rate, miscarriage rate, multiple pregnancy rate and pregnancy outcome.

Nine randomized controlled trials comparing IVF with and IVF without PGS were included in our meta-analysis. Fluorescence in situ hybridization (FISH) 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% CI -0.13 to -0.03). For a live birth rate of 26 percent 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.

We concluded that 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.

This inefficacy of PGS using cleavage stage biopsy and FISH, led in recent years to a renewed and increasing interest in further development of, and adjustments to, the PGS technique. New methods to determine the ploidy status of a single cell are in development, such as comparative genomic hybridization (CGH) arrays or single nucleotide polymorphism (SNP) arrays (Wells et al., 2008; Hellani et al., 2008; Vanneste et al., 2009; Schoolcraft et al., 2009b;
Treff et al., 2010; Gutierrez-Mateo et al., 2011). In an attempt to avoid the confounding effects of chromosomal mosaicism, embryos are now biopsied at either the zygote or blastocyst stage (Jansen et al., 2008; Geraedts et al., 2009). In addition, increasing time and money is invested in the development of high-tech non-invasive methods to select the best embryo for transfer in IVF. This includes metabolomic profiling, amino acid profiling, respiration-rate measurement, and birefringence imaging (Nagy, 2008).

All research on the development and optimization of embryo selection methods is driven by the concept that embryo selection is essential to optimize the success rates of IVF, since embryos that are cryopreserved have a reduced chance of implanting after thawing. Better selection methods should thus result in higher live birth rates without an increase in multiple pregnancies.

Recent developments challenge this concept. We discussed these developments and hypothesized that the path of embryo selection is turning into a dead-end in the quest for optimal IVF success rates (Chapter 7).

The main reason is the accumulating evidence that all embryos can now be cryopreserved and transferred in subsequent cycles without impairing pregnancy rates or maybe even with an improvement in pregnancy rates (Aflatoonian et al., 2010). In such a scenario no selection method will ever lead to improved live birth rates, as, by definition, the live birth rate per stimulated IVF cycle can never be improved when all embryos are serially transferred. In fact, if the selection method under study would not be 100% specific, this would then only lower the live birth rate after IVF. The only parameter that could possibly be improved by embryo selection would be time to pregnancy, if embryos with the highest implantation potential are transferred first.

Based on the data and thoughts summarized in this thesis, PGS using cleavage stage biopsy and FISH should no longer be offered to women as a means to increase live birth rates after IVF.

New forms of PGS are currently under development. All are at the stage of method assessment or at the stage of pilot studies and no trials that investigate the value of these new approaches in PGS have yet been registered. Nonetheless, these methods have, at least in some clinics, already been implemented in clinical practice. Indeed, the theory behind these developments sounds plausible, the techniques are attractive and provide potential, and the first results seem promising (Hellani et al., 2008; Schoolcraft et al., 2009b; Treff et al., 2010; Gutierrez-Mateo et al., 2011). However, a decade ago, exactly the same was said when PGS using cleavage stage biopsy and FISH was developed and introduced into clinical practice.

Any fair-minded person will now conclude from this thesis that for many women PGS has been a significant additional cost without robust evidence of any benefit. For many women their chance of a live birth has been harmed. It is important to note that many issues that are suggested in this thesis to undermine PGS, 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 first provide sufficient evidence that they work, and this should then be followed by rigorous RCTs to provide evidence for the efficacy of these new techniques.

In case upcoming studies indeed show that all embryos of an IVF cycle can be cryopreserved and transferred in subsequent cycles without impairing pregnancy rates or maybe even with an improvement in pregnancy rates, then such RCTs are of limited value, as no form of PGS will ever improve the live birth rate after IVF in such a scenario, but will only be able to potentially decrease time to pregnancy.

Future research on optimizing IVF success rates should focus less on the development and optimization of embryo selection techniques but more on the optimization, evaluation and broad implementation of the best cryopreservation protocols. Decision models need to be developed to determine the most cost-effective transfer policies of cryopreserved embryos.

Potential benefit can also be expected from a more widespread use of proper methodology in studies evaluating methods, materials, and techniques in the IVF-laboratory. More rigorously designed and executed randomized controlled trials, with proper power calculations, using the right outcome measures, are needed as they will ultimately lead to better ‘evidence based laboratory practice’ and better IVF-results.

From a biological perspective, further studies are needed to elaborate on the origin of chromosomal mosaicism in human preimplantation embryos, on factors that affect its incidence, as well as on the fate of diploid-aneuploid mosaic embryos.

References


Centers for Disease Control and Prevention, American Society for Reproductive Medicine, and Society for Assisted Reproductive Technology (2010) 2008 Assisted Reproductive Technology Success Rates: National Summary and Fertility Clinic Reports.


Harper, J., Sermon, K., Geraedts, J., Vesela, K., Harton, G., Thornhill, A., Pehlivan, T., Fiorentino,


