European Journal of Human Genetics (2009) 17, 1550–1559

www.nature.com/ejhg

ARTICLE

Preimplantation genetic diagnosis for mitochondrial DNA disorders: ethical guidance for clinical practice

Annelien Bredenoord*[,1,2], Wybo Dondorp[1,2], Guido Pennings[3], Christine de Die-Smulders[4], Bert Smeets[5] and Guido de Wert[1,2]

[1]*Department of Health, Ethics and Society, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands;* [2]*Research Institute Growth and Development GROW, Maastricht University, Maastricht, The Netherlands;* [3]*Department of Philosophy and Moral Sciences, Bioethics Institute Ghent, Ghent University, Ghent, Belgium;* [4]*Department of Clinical Genetics, Maastricht University Medical Centre, Maastricht, The Netherlands;* [5]*Department of Genetics and Cell Biology, Maastricht University Medical Centre, Maastricht, The Netherlands*

Although morally acceptable in theory, preimplantation genetic diagnosis (PGD) for mitochondrial DNA (mtDNA) disorders raises several ethical questions in clinical practice. This paper discusses the major conditions for good clinical practice. Our starting point is that PGD for mtDNA mutations should as far as possible be embedded in a scientific research protocol. For every clinical application of PGD for mtDNA disorders, it is not only important to avoid a 'high risk of serious harm' to the future child, but also to consider to what extent it would be possible, desirable and proportional to try to reduce the health risks and minimize harm. The first issue we discuss is oocyte sampling, which may point out whether PGD is feasible for a specific couple. The second issue is whether one blastomere represents the genetic composition of the embryo as a whole – and how this could (or should) be investigated. The third issue regards the cutoff points below which embryos are considered to be eligible for transfer. We scrutinize how to determine these cutoff points and how to use these cutoff points in clinical practice – for example, when parents ask to take more or less risks. The fourth issue regards the number of cycles that can (or should) justifiably be carried out to find the best possible embryo. Fifth, we discuss whether follow-up studies should be conducted, particularly the genetic testing of children born after IVF/PGD. Finally, we offer the main information that is required to obtain a truly informed consent.
European Journal of Human Genetics (2009) **17,** 1550–1559; doi:10.1038/ejhg.2009.88; published online 27 May 2009

Keywords: ethics; PGD; mitochondrial DNA; genetic testing minors

Introduction

When applying preimplantation genetic diagnosis (PGD) for mitochondrial DNA (mtDNA) disorders, interpretation of the test results may be difficult and it is conceivable that in some cycles only affected embryos are available for transfer. Instead of enabling parents to have a healthy child, PGD then may only contribute to a reduction of reproductive risk by selecting the embryos with the highest probability of leading to a healthy child. Earlier,[1,2] we discussed the pros and cons of PGD for mtDNA disorders, concluding that despite the drawbacks no convincing moral arguments exist to regard risk-reducing PGD as unacceptable. Nevertheless, several ethical questions arise

*Correspondence: A Bredenoord, Department of Health, Ethics and Society, Faculty of Health, Medicine and Life Sciences, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands.
Tel +31 641 074 365; Fax: +31 433 670 932;
E-mail: a.l.bredenoord@umcutrecht.nl
Received 25 September 2008; revised 18 March 2009; accepted 23 April 2009; published online 27 May 2009

when considering application in clinical practice. These specific ethical questions are generated by the characteristics of mtDNA genetics (such as heteroplasmy, which is a mixture of normal and mutant mtDNA, the level of which can differ among tissues). When a couple asks for PGD for a mtDNA mutation, is it morally justified to honor their request and to (experimentally) offer them PGD? Or is this still premature, taking into account the risks and uncertainties? After all, the reasonable welfare standard, which we consider to be the best way to take into account the welfare of the child, allows PGD for mtDNA mutations insofar, as there is no 'high risk of serious harm' to the future child. This standard entails the view that for assisted reproduction to be justified, the child to be must have a reasonable chance of an acceptable quality of life.[2–4] What are the implications of this standard for determining the heteroplasmy levels (the cutoff points) below which embryos are considered to be eligible for transfer? Does sufficient knowledge exist about whether the analyzed cell(s) represents the entire embryo? And to what extent should one search for better embryos, for example by starting another cycle of *in vitro* fertilization (IVF)?

In this paper, we will examine the questions that may successively arise when PGD for mtDNA disorders is considered for application in clinical practice. The central question of this paper is: which moral requirements have to be met before applying PGD for mtDNA disorders in clinical practice? We will not offer a clear-cut recipe, but we aim to present the ingredients, the points to consider when offering PGD for mtDNA mutations.

Three further preliminary remarks should be made. Firstly, little experience exists with PGD for mtDNA mutations (thus far only two applications have been reported).[5,6] Some considerations may therefore be to some extent speculative. This is unavoidable if one aims to practice ethics pro-actively and this is a reason why we discuss several scenarios. Secondly, we will in this paper only consider blastomere biopsy, as PGD on blastomeres currently has a higher efficiency and accuracy than the other approaches.[7,8] Thirdly, if PGD would indeed be offered, good reasons exist to embed the first applications of PGD for mtDNA mutations in a scientific research protocol. This is the most efficient way to obtain systematic data and to increase knowledge in mtDNA genetics. A further advantage is that it guarantees that patients are sufficiently aware that they are enrolled in an experimental treatment. Obviously, the rarity of some mtDNA mutations may limit the feasibility of clinical trials. In such cases, efforts should be made to conduct (international) multi-centre trials.

Three categories of mutations

Most ethical difficulties in mitochondrial clinical practice are related to interpretation and segregation of heteroplasmy. PGD for homoplasmic mutations (that is, only mutant mtDNA is present) such as Leber Hereditary Optic Neuropathy raises different types of ethical questions.[2] This paper is therefore restricted to the heteroplasmic mutations. The reliability of PGD for those mutations will depend on whether the specific mutation shows:[2,9]

1. a close correlation between the mutant load (that is, the ratio of mutant to normal mtDNA) and disease severity;
2. a uniform distribution of mutant mtDNA in all blastomeres;
3. no change in mutant load with time (both prenatally and postnatally).

We will in this paper refer to three categories of mutations.

Category I regards stable mutations with a predictable outcome. These mutations fulfill the three criteria above.[10–12] The main examples are the mutations *m.8993T>G* and *m.8993T>C*, leading to the neurodegenerative diseases NARP (Neurogenic muscle weakness, Ataxia, Retinis Pigmentosa) and Leigh syndrome. Both mutations have a strong genotype–phenotype correlation and show very little tissue-dependent or age-dependent variation in mutant load.[10–12] For these mutations, in general, symptoms seem to occur above ~50% mutant load and the frequency and severity increases with increasing mutant load.[10] Patients with severe NARP syndrome generally have mutant loads above 70–80%, although some exceptions have been reported as well.[10] Patients with severe Leigh syndrome generally have mutant loads above 80%.[10] Nevertheless, a twofold cautiousness is required. Firstly, cautiousness regarding the exact cutoff levels, as statistical variation of the mutation percentages is rarely supplied and depends highly on the technology used and the number of patients tested. Secondly, cautiousness regarding the interpretation of 'seriousness'. Although some general outlines may be given, what exactly constitutes 'serious symptoms' is something to be assessed case-by-case, depending on the specific circumstances. Although the *m.8993T>C* mutation (Leigh) is generally considered to be clinically milder than the *m.8993T>G* mutation (NARP) and the threshold for disease expression seems to be slightly higher, we discuss these mutations here together as the nucleotide *8993* mutations. We presume that also during embryonic and fetal development, tissue and time-dependent variation in mutant load is very limited. Above 80%, the child is very likely to be affected. There is a gray zone between 50 and 80% mutant load, especially for the *m.8993T>G* NARP mutation.[11,13,14] This zone is 'gray' in the sense that there is a chance of developing symptoms, but the frequency and seriousness are not exactly known, and in the sense that precise numbers are lacking because of technical variables. One may reasonably assume that an embryo with

a mutant load below 50% is very likely to result in a healthy child.

Category II are mutations, which are unstable in time with a non-uniform tissue distribution and with no reliable genotype–phenotype prediction on the basis of mutant load. They do not fulfill the criteria above.[9,15] An example of this is the *m.3243A>G* mutation leading to MELAS, which is one of the most common mtDNA mutations. The clinical expression of this mutation is very variable, the classical presentation being the combination of mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes, whereas in other families only diabetes mellitus and deafness are found. However, not only interfamilial, but also intrafamilial, variation may be huge and is only partly related to mutation load. In general, symptoms may occur above 30% and the frequency and severity increases with increasing mutation load.[16] However, exceptions occur and prognostic predictions on the basis of mutation load are not reliable for individuals.[15]

Category III contains mutations, which are private or family-specific point mutations and for which no additional information is available. Insufficient evidence exists to decide whether they fulfill the criteria mentioned and allow reliable PGD.

Feasibility and representativity

Feasibility: oocyte sampling?

Oocyte sampling is an option to be considered to answer the question whether IVF and PGD will be feasible (that is, would have a fair chance of success) for a particular couple. After a superovulation cycle, all the oocytes of the woman (on average 10–12 oocytes after ovulation induction) are taken and analyzed to assess the chance that eventually embryos with no or a low-mutant load are available.[17] For the *8993* mutations, an atypical (skewed) segregation of mutant load among the oocytes may be expected.[6,18] This would be favorable, because this increases the chances of obtaining low or zero-mutant embryos (although the chance of obtaining high mutant load embryos will be increased as well). For the *m.3243A>G* mutation, a more continuous distribution of mutant load may be expected.[19] When it would become clear beforehand that the chance of obtaining oocytes with no or a very low-mutant load is negligible, there seems to be no use in starting an IVF/PGD procedure. For those ascribing high moral value to embryos, an advantage of oocyte sampling is that unnecessary creation and destruction of embryos is avoided. Another advantage is the (material and immaterial) costs that may be saved, as there will be no fertilization and further PGD analysis (and subsequently no complex decision making regarding which embryo to transfer). However, these advantages only apply for women with oocytes containing high mutant loads. A major disadvantage for women for whom PGD does seem to be feasible is that they have to undergo a novel cycle; their analyzed oocytes are lost and thus cannot be used to establish a pregnancy. Therefore, we propose to sidestep oocyte sampling and proceed directly to IVF and PGD.

Representativity: a uniform distribution?

Before PGD can be considered for an mtDNA mutation, it is important to ascertain that the proportional levels of mutant to wild-type mtDNA quantified in the biopsied cell are representative of the levels in the embryo as a whole.[8] This issue of representativity is of utmost importance from a moral point of view: would there not be a uniform distribution, then the cell(s) taken for analysis do(es) not represent the genetic composition of the whole embryo. This may result in the birth of a severely affected child. Data of neutral polymorphisms in mice and human, and data regarding the *8993* mutations indicate that the intercellular variation of mitochondrial heteroplasmy is fairly small among most preimplantation embryos.[6,8] Data are limited for the *m.3243A>G* mutation. Blastomere representativity could be investigated at two stages: preclinically (IVF/PGD without the intention of embryo transfer) and clinically (IVF/PGD with the intention of embryo transfer). What would be the best route to ascertain this?

Preclinical route In a preclinical route, embryos carrying an mtDNA mutation could be created and dissected to find out whether all blastomeres carry the same amount of mutant mtDNA. If it turns out to be the case, this can for the moment be taken as sufficient proof of blastomere representativity for the mutation in question. The only possibility to obtain embryos carrying a specific mtDNA mutation would be to create them by means of IVF. This requires donation of gametes from both a male and a female donor, whereby the oocyte donor should carry the mtDNA mutation in question. This leads, firstly, to an ethical assessment of oocyte donation for research purposes and, secondly, to an ethical assessment of creating embryos for research purposes. Given the burdens and risks of the procedure, oocyte donation is morally sensitive. We, nevertheless, consider it to be morally acceptable, provided the risks are minimal, the donor is sufficiently informed about the risks and discomforts, and voluntariness is safeguarded.[20,21] In most countries creating embryos specifically for research is both legally prohibited and morally rejected, particularly because of the contested moral status of the embryo. Many would only accept the creation of embryos in the context of a parental project. But what would be wrong with creating embryos in the context of preparing for a possible parental project? The final goal would be to enable couples at risk to conceive healthy children. Although those first embryos will not be used directly for a pregnancy, subsequent embryos will be used, provided the results are encouraging. Although in some

cases an extra step is necessary, the ultimate goal remains the conception of children. It cannot be maintained that this is morally very different from the current IVF practice, with its implied acceptance of creating and discarding surplus embryos.[22] Once one accepts the creation and killing of embryos to benefit infertile people with a child-wish, it seems questionable to argue that the creation and killing of embryos for research (which eventually may benefit ill people) is condemnable.[23,24]

Clinical route Another way of obtaining more evidence for embryo representativity could be at the beginning of a clinical application of IVF/PGD. Two approaches are conceivable. A first approach may be to start an IVF-cycle and to examine two cells of each embryo available (usually 10–12 embryos are available). This may sufficiently prove whether there is a uniform distribution of mutant mtDNA in all blastomeres. If there would be a problem with an uneven segregation, that is, large differences in mutant load, this would probably become clear (although this cannot be stated with certainty, as there will always be a limited number of samples). However, two-cell biopsy is controversial, as current data indicate that it may impede success rates.[25–27] Cohen *et al* argue that 'this (two-cell biopsy) approach should be reserved for cases in which diagnostic accuracy is considered paramount, more important than embryo implantation'.[25] With regard to the mtDNA mutations, diagnostic accuracy actually is more important than implantation. A two-cell biopsy therefore would be justifiable, at least in the first applications: if the mutant load can be reliably determined, one can return to single cell biopsy, and if this is not the case, one should stop offering PGD.

A second approach would be to test a single cell of each embryo. Subsequently, all remaining blastomeres of those embryos with a high mutant load are analyzed within 24 h (to be able to use this information before implantation). If it turns out that the mutant load for all tested embryos is similar in all blastomeres, this indicates that one blastomere represents the genetic composition of the embryo as a whole. This second approach may be conducted to avoid a two-cell biopsy. It can also be used to determine whether one or two blastomeres should be used for the second cycle.

Both preclinical and clinical investigation may thus indicate whether PGD is reliable. What would be the best approach? If the 'clinical' route indeed provides sufficient evidence, then this route is preferable over the preclinical route. After all, the most important goal – assuring that the genetic test is reliable – can be met and a pregnancy would still be within the possibilities. This fits both the interests of the woman (that is, a healthy pregnancy) and the principle of proportionality (that is, the risks, costs and burdens of a procedure balanced against the expected outcome). As the clinical route is likely to consume fewer embryos than the preclinical one, the clinical route is also preferable from an embryo-saving perspective.

Determination of cutoff points

A cutoff point is the threshold of mutant load above which no embryos are considered for transfer. If only embryos above the threshold are found, further options are limited to either trying again by starting a new IVF/PGD cycle or deciding to stop trying PGD. It is important to stress that a cutoff point figures as the lower limit. It points out what amount of risk is still considered acceptable. Searching for embryos with a lower mutant load is not morally neutral. When opportunities exist to improve the outcome, that is what one should aim at – taking into account other morally relevant aspects as well. This implies that for every clinical application of PGD for mtDNA disorders, it is not only important to be sure that helping this specific couple would be justifiable in view of the reasonable welfare standard, but also to consider to what extent it would be possible, desirable and proportional to try to further reduce the health risks by adapting procedures or trying another cycle. Of course, these are questions both for the medical team and the couple.

For all mtDNA mutations, health risks could be reduced substantially by determining a low cutoff point. However, the lower the cutoff points are fixed, the more embryos are discarded, and thus, the lower the 'take home baby rate'. On the other hand, the higher the cutoff points are fixed, the more embryos will probably be available. The price to be paid is an increased risk that the welfare of the child will be compromised. In the process of determining the appropriate cutoff point, both the proportionality and the welfare of the child need to be assessed. After all, the risks and burdens (material and immaterial) of IVF and PGD must stand in a reasonable proportion to the desired aim, which is the conception of a (healthy) child.

What, then, are the appropriate cutoff points? For the *8993* mutations (category I), levels of mutant load below 50% seem compatible with a healthy phenotype.[10,11,28] A gray zone exists between 50 and 80%. Whether the reasonable welfare standard also allows embryos in the gray zone to be transferred is debatable and also depends on the specific mutation. The more the mutant load is heading towards the 80%, the higher the probability of (severe) NARP or Leigh symptoms. Nevertheless, a cutoff point of 80% may be justifiable as well, bearing in mind that a cutoff point is the absolute lower limit. This implies that if one takes a cutoff point of 80%, physicians and parents should consider to make efforts to search for embryos with a lower mutant load.

For the *m.3243A>G* mutation (category II), the threshold to disease expression is not exactly known and can vary among individuals, which makes determining the cutoff point an intricate enterprise. As a mutant load below 15% does not seem to result in a severe phenotype, it seems in accordance

with the reasonable welfare standard to determine the cutoff point for the *m.3243A>G* mutation at 15%.[16,29] Although also here a certain margin exists, embryos carrying 15% mutant load probably do not run a high risk of serious harm.

Characteristic for the group of private mtDNA mutations (category III) is that they are only found incidentally, in only a few families or even a single one, and that therefore insufficient information is available to judge if these mutations allow sufficiently reliable predictions. If more information about a specific mutation would be available, then that mutation would not fit this category any longer (it would shift to category I or II). PGD for any point mutation may be offered as long as mutant-free embryos are available and can be reliably identified for transfer. The problem of transferring embryos with some mutant load is that it is difficult to determine a cutoff point for each private mutation because of the lack of data. For some private mutations, a cutoff point may be derived through extrapolation of what is known about other mutations, as for example has been done for the *T9176C* mutation.[30] Another option is to determine the biochemical cutoff point in a cell line system (cybrids). The rationale is that clinical manifestations are unlikely to occur if the mutation load does not lead to an enzyme deficiency. Nevertheless, as this is rather speculative, we will restrict ourselves here to category I and II, the *8993* and *m.3243A>G* mutations (which raises similar ethical questions as the private mutations).

The proposed cutoff points above should be regarded as provisional. They should be adjusted when more experience exists with PGD for mtDNA mutations, or when more is known about the threshold to disease expression (and for some mutations complete clarity may never be obtained because of the clinical variability and the possible changes in mutant load).

Taking more risks?

It is conceivable that a couple does not agree with the proposed cutoff point (which may already become clear in the pre-test counseling or during the procedure). Although the initial aim of a couple may have been to look for an embryo below the cutoff point, no suitable embryos may become available during the IVF/PGD procedure. It is conceivable that couples with a very strong wish for a child are satisfied with the transfer of an embryo with a mutant load above the determined cutoff point. This may be especially the case for couples depending on IVF anyway because of fertility problems (which broadly speaking concerns 50% of PGD users[31]). Is it acceptable, then, to follow their wish? A (slightly) flexible use of the cutoff points may be defended on grounds of two considerations.

The first consideration is whether the acceptability of taking more risk depends on the number of cycles already carried out. One can imagine that a couple would be more inclined to ask for the transfer of embryos with (some) mutant load in a third or fourth cycle than when they just have started IVF/PGD. Although it is a request easily sympathized with, adhering to the reasonable welfare standard implies that a high risk of serious harm for the resulting child should be avoided in any case. The transfer of an embryo with a mutant load that is likely to cause a severe phenotype is morally unjustified. The second consideration is that some arbitrariness regarding the cutoff values cannot be avoided and that the safety margin taken can be debated, especially for couples in whose family the mutation seems to show a mild phenotype (eg, diabetes mellitus in case of the *m.3243A>G* mutation). How far do you need to be on the safe side? Which arguments are valid to adjust the cutoff point? Any sensible argument either directed on the percentage as such, the clinical interpretation or the safety margin should be discussed on a case-by-case basis by the health care professionals involved in PGD.

There may also be couples who want to take less risk or no risk at all. They, for example, only regard mutant-free embryos as eligible for transfer. This will be examined below, when we discuss the number of cycles.

Another situation that may occur is that a choice has to be made between a low mutant load embryo that is morphologically of lesser quality and a higher mutant load embryo that is morphologically of better quality. Here, one has to maneuver between the futility of transferring an embryo of very poor morphology on the one hand and not exceeding the cutoff point on the other hand.

Number of cycles

Chances of eliminating (or at least minimizing) health risks may increase with the number of embryos. The more embryos available, the higher the chances of obtaining embryos with zero or low(er) mutant load. The number of embryos can be increased in two ways. The first involves a stronger ovarian stimulation. This could provide more oocytes and thus more embryos. Whether it is acceptable to offer a more aggressive hormone treatment to obtain more embryos has already been a point of debate with regard to PGD for translocations.[32] It is quite contrary to the current tendency to mild ovarian stimulation.[33] Nevertheless, it should not be completely ruled out, especially when the woman belongs to a low-risk group for developing ovarian hyperstimulation syndrome.

A second way to obtain more embryos is by conducting more cycles. Suppose that IVF/PGD for the *m.3243A>G* mutation results in one embryo with a mutant load of 12%. Or, when PGD is carried out for one of the *8993* mutations, an embryo with 70% mutant load is available (being in the so-called gray zone). Is it morally justified to transfer these embryos? Or should the medical team propose to start a new cycle to find a 'better' embryo?

Should one start a new cycle?

In other applications of PGD, the medical team will not propose a new cycle if suitable embryos are at hand. So why would they do so here? Although one should, other things being equal, strive to minimize risks, this should also be weighed against the implications. In view of the (physical, emotional and financial) burdens of both IVF and PGD (both for the couple and for society), it can be questioned whether conducting a new cycle is justified. Many couples experience an IVF/PGD cycle as extremely stressful.[34] Furthermore, cryopreservation of biopsied embryos is not very successful, at least not at this moment (eg, Magli MC *et al*[35] and Parriego M *et al*[36]). The chance exists that only embryos in worse condition are obtained when one starts a second or third IVF/PGD cycle. The overall result then would be that suitable embryos were available in an earlier round but that these cannot be used anymore because they did not survive thawing. To start a new round when a suitable embryo is available is thus quite a gamble. So is this still proportional?

A reply to this may be that the proposed cutoff points are debatable at the state of knowledge. One cannot say with certainty that an embryo in the gray zone (*8993* mutations) or an embryo with, for example, 12% mutant load (*m.3243A>G* mutation) will result in an unaffected child. One could therefore prefer to obtain embryos with the lowest possible mutant load as long as this application of PGD is still novel and experimental. Starting a new cycle therefore is morally acceptable.

Moreover, it is shown for at least a few families that obtaining embryos without mutant load is realistic.[6] One could therefore argue that if one chooses to use a risky technology and a possibility exists to considerably reduce the risks, then physicians and the couple should at least consider to do so. One could in this light even consider to preferentially transfer male embryos to reduce the risk of transmission to next generations.

In view of our discussion above, we now may delineate a rough framework about when (not) to start another IVF/PGD cycle and which embryo(s) to transfer (we accept the general consensus regarding the transfer of one, maximally two embryos). Although it seems best in theory to look for 'the best possible embryo', in practice this may be no sinecure. The lower limit in all cases is the avoidance of a high risk of serious harm for the child. More concrete, on the current state of knowledge we have proposed to use a cutoff point of 15% for the *m.3243A>G* MELAS mutation. For the *8993* NARP/Leigh mutations, we have proposed to use a mutant load of 80% as the cutoff point. For many private mutations, it is impossible because of the lack of data to offer a well-founded cutoff point. For those mutations, PGD can be offered but only embryos with zero-mutant load seem eligible for transfer. Although it is morally acceptable that a couple and the medical team confer about whether or not to start a new cycle, looking for better embryos should not be seen as morally required if an embryo below the cutoff point is available. The precise number of cycles is something that should be determined on a case-by-case base, also depending on the specific mutation, the wishes of a couple, their individual chances of success, the number of effective transfers and the number of cycles allowed and reimbursed in a country.

Scientific research after IVF/PGD

After the IVF/PGD procedure, not all uncertainties surrounding PGD for mtDNA mutations will be resolved. This will particularly be the case for the unstable mutations with an unpredictable outcome (that is, category II, such as the *m.3243A>G* mutation). Residual uncertainties mainly regard whether the mutant load has changed in time and whether there is a uniform distribution of mutant mtDNA in all blastomeres. Various types of post-procedure scientific research may contribute to the further clarification of these uncertainties.

The first type of scientific research concerns analysis of the embryos that will not be transferred. Probably, the most interesting aim of this research is to check the segregation of mutant load between the blastomeres. To the extent that this would provide an extra safety check relevant for further applications, this type of research is recommendable. It is desirable that the medical team asks the parents to donate their affected embryos for scientific research (as is standard procedure in other applications of PGD).

Another option to check for reliability of the PGD test regards prenatal diagnosis (PND), using CVS or amniocentesis during the pregnancy. However, as we cannot do justice to the complex issue of PND for mtDNA disorders within the scope of this paper, and as debate is possible about the added value of PND for mtDNA disorders, we will not further discuss this option. Further interdisciplinary debate on whether (and if so, under what conditions) PND should be offered after PGD is desirable.

Testing children

We will focus on a third type of scientific research, that is, the genetic testing of the children born as a result of IVF/PGD for an mtDNA mutation. The main goal would be to validate the PGD procedure by ascertaining that the genetic diagnosis made at the embryonic stage is a reliable predictor of the health of the resulting child. Clearly, periodic clinical examinations could also add to this aim, but only after years and not as clearly as a genetic test could do. The danger is that PGD for mtDNA mutations would be introduced into the clinic and would turn out to be unreliable only after years (when the first children develop disease symptoms). The added value of a genetic test is that it could make clear immediately

whether the percentage of the mutation is constant from the embryonic stage throughout pregnancy until birth and thereafter. If not, a reliable prediction cannot be made. One should then either re-assess the criteria for embryo selection or refrain from a next application of PGD. Therefore, the importance of the knowledge possibly generated by genetic testing seems beyond question, but several ethical issues need to be scrutinized before one considers testing a newborn.

If the newborn will be genetically tested, three scenarios are possible. In a first scenario, the postnatal test confirms the genetic diagnosis obtained in the embryonic phase. This suggests that PGD is reliable. In a second scenario, the mutant load has decreased. Although this is favorable for the child, strong doubts will arise regarding the reliability of PGD. In a third scenario, the postnatal test shows a (considerably) higher mutant load. Not only strong doubts will arise regarding the reliability of PGD, this also results in adverse predictive information about the future health of the (possibly asymptomatic) child. Generally speaking, predictive genetic testing of minors is only considered ethically acceptable provided that it is to the medical benefit of the child. In our case, the child undergoing follow-up would probably not medically benefit from testing, although lack of clarity regarding the therapeutic value of early detection exists. Some may argue that lack of cure does not equate lack of treatment.[37] Regular check-ups and some preventive health measures may be beneficial for the child. Others may reply that this is often also done just 'to do something'. Furthermore, although some argue that children may psychosocially benefit from knowledge about their genetic status,[38,39] the majority still view that children should not be tested for untreatable late onset disorders (see Bredenoord AL *et al*[2]). If we assume that follow-up studies are to be classified as non-therapeutic medical research with incompetent participants, what, then, about the ethics?

A research procedure which is not intended directly to benefit the participating child is not necessarily unethical as long as it meets the following requirements.[40–42] Firstly, a preliminary requirement is that the research satisfies the material and procedural requirements for scientifically and ethically sound medical research. This requirement constitutes an additional argument to embed – if feasible – the first applications of PGD for mtDNA mutations in a scientific research protocol. Secondly, 'these individuals (incompetent research subjects) must not be included in a research study that has no likelihood of benefit for them unless it is intended to promote the health of the population represented by the potential subject' (WMA, 2008 art 27). The results of the intended research should mainly benefit the group of patients to which the participant belongs. Whether this requirement can be met depends on the interpretation of 'the population represented'. If this phrase should be understood as referring here (more broadly) to regards all the people carrying mtDNA mutations, then this requirement can be met. After all, the development of reliable preimplantation testing will benefit those carrying mtDNA mutations by giving them reproductive options. Therefore, it promotes their reproductive health. Furthermore, the increased knowledge of mtDNA genetics may eventually contribute to the development of treatment as well, but this is highly speculative. If, however, 'the population represented' only includes minors carrying an mtDNA mutation, then it is questionable whether this requirement is met. Thirdly, 'the research cannot instead be performed with competent persons' (WMA, 2008 art 27). This is the case, as the research can only be conducted on children born after PGD and/or PND for an mtDNA disorder. To wait until the child has reached adulthood is not a real alternative. Fourthly, the health and safety risks and the burdens for the research participant should be 'minimal'. Whether the follow-up studies can meet this last requirement depends on different variables in the research design.

A first factor is which tissue needs to be tested. If the use of existing material (eg, hair, urine and cord blood) could do the job of detecting the mtDNA mutation,[43] then a muscle biopsy would be unnecessary and unjustified. A second factor is the timing of testing: is it sufficient to collect the material directly after birth, or are parents asked to come back later? Related to this is a third factor, namely the frequency of testing. A fourth factor is whether the test results will be disclosed to the parents.

This latter issue is complex. As the data are easily retraceable to the child in question, anonymity is not possible. In theory, three options therefore remain: (1) to refrain from genetic testing, (2) genetic testing with disclosure to the parents and (3) genetic testing with non-disclosure (that is, withholding the test results from the parents). In practice, the feasibility of withholding test results from the parents is questionable. Therefore, the dilemma is as following. If one refrains from genetic testing, the reliability of PGD cannot be ascertained (or only after years). If one discloses the test results, the predictive genetic information could (psychosocially) harm the child. But, if feasible at all, when one withholds the test results, parents are not able to use the information for possible extension of their family. If it turns out that the PGD procedure did not prevent the birth of a child with a high mutant load, this may guide them and others in future reproductive decisions. If, to the contrary, the child indeed carries no or a low mutant load, they and future users may gain reproductive confidence.

The main ethical assessment to be made, then, is whether disclosure of the knowledge generated by the test results violates the fourth criterion that states that only minimal risk and burden is acceptable. If the child turns out to have a (much) higher mutant load than expected, this may burden the child in a double sense. Firstly, the

child grows up with knowledge about its adverse health prospects. How much weight we should give to this knowledge also depends on the expected seriousness of the disease, the age of onset and the chance that symptoms will develop. As only embryos with low mutant load will be transferred, it might be a reasonable guess to estimate the chance that the child will develop severe disease symptoms as fairly low. The problem is that this cannot be stated with certainty and that this is precisely the information that the research aims to clarify. A mitigating factor may be that the genotype of the child is already known, albeit in the embryonic phase (and perhaps also in the fetal phase, if a PND has been carried out). The child is born as a result of PGD: it will know about the presence of an mtDNA mutation in the family anyway, also when it is not tested. The second drawback of disclosure may be that a thus far symptom-free child is heading towards a life of hospital visits, check-ups and so on, although no real treatment is available (but the child may have periodic clinical examinations anyway). As we explained above, opinions will differ on how much weight we should ascribe to these drawbacks.

Involving children in this type of research is morally complex. A tension exists between the need to ensure that PGD is safe, effective and reliable and the need to protect the participating children (eg, Wendler D and Varma S[44]). This tension also exists in pediatric drug research. The importance of clinical trials with children *versus* the acceptability of exposing children to risk has been a topic of increased debate. The question is whether these requirements can be reconciled, or whether the minimal harm standard needs adjustment (eg, Wendler D and Varma S[44] and Caldwell PHY[45]). In our case, the risks involve psychosocial harm as a result of knowledge about an unexpected high mutant load. We estimate the probability that the child will experience such harm to be small, but real. It is questionable whether these risks satisfy a strict interpretation of the minimal harm standard. If a more flexible interpretation of the minimal harm standard is considered to be acceptable (in the United States, for example, one accepts a minor increase over minimal risk[40]), and if the requirement that this research is necessary to promote the health of the population represented is interpreted flexibly, then these follow-up studies may be ethically justifiable. Even more so because the child will know about the genetic risk in the family anyway.

Although being aware of the many ethical pitfalls, we are inclined to recommend conducting follow-up studies in the context of mtDNA disease. In close deliberation with the parents one could decide whether and how to disclose the test results. Clearly, this complex topic needs further interdisciplinary debate. This debate should regard both the minimal harm standard, the interpretation of the 'population represented' and the acceptability of predictive genetic testing of children.

Table 1 Information required to support informed consent

Scientific research
Ensure that the couple is aware that PGD for mtDNA mutations is not regular practice but experimental (in case PGD is embedded in scientific research).
Residual risks
Discuss with the couple that PGD may only reduce reproductive risk; they are sufficiently aware that residual health risks may be unavoidable.
Consider to what extent it is possible, desirable and proportional to try to further reduce the health risks.
Moral minimum: avoid a high risk of serious harm
Alternatives
Discuss possible alternatives such as PND, oocyte donation and adoption. The couple has weighed the pros and cons of these strategies.
Feasibility
Bring up the possibility (and pros and cons) of oocyte sampling to check the feasibility of PGD.
Reliability
Ensure that the genetic test result is sufficiently representative. This is to be clarified before the IVF/PGD procedure will be carried out further. Ensure that the couple is aware that the IVF/PGD procedure will only be continued when the mutation load can be reliably detected and is representative.
Cutoff point
Discuss the determined cutoff point.
Discuss whether priority should be given to embryos of good morphology or embryos with a lower mutant load.
Number of cycles
Discuss and determine the maximum number of cycles.
Scientific research after IVF/PGD
Insist on the possibility to donate affected embryos for scientific research.
Discuss the possibility of regular clinical examination.
Discuss with the couple the importance and pitfalls of conducting a genetic test after birth; if they decide to enroll, they should have made a decision about whether or not to disclose the test results (and be aware of the possible implications).

Informed consent

Obtaining an adequate informed consent is one of the most important moral requirements. What constitutes an 'adequate' informed consent in our discussion? What should be the ingredients to be discussed during the counseling process? Table 1 lists the issues that the counselor should bring up for discussion with the prospective parents.

Obviously, guiding the couple to make a truly well-informed decision requires both time and professional counseling skills. In the end, much hinges on a transparent communication between prospective parents and the medical team.

Conclusion

When considering PGD for an mtDNA mutation, PGD may only contribute to a reduction of reproductive risk. Although morally acceptable,[2] both the couple and the medical team should carefully look for possibilities to minimize the risks and harms as much as possible. In this paper, we offered the points to consider when applying PGD for mtDNA mutations in clinical practice. We have discussed the questions that may arise in clinical practice on the basis of three categories of mtDNA mutations, but this framework of course also applies for other heteroplasmic mtDNA mutations. Although far from ideal and with considerable drawbacks, we think that PGD may be a valuable reproductive option for couples at risk of transmitting an mtDNA mutation to their offspring.

Acknowledgements

This work was supported by a grant from the European Union Sixth Research Framework Programme (the MITOCIRCLE project contract no 005260). We thank Professor Dr J Geraedts for his valuable comments.

References

1 Bredenoord AL, Pennings G, Smeets HJ, De Wert GM: Dealing with uncertainties: ethics of prenatal diagnosis and preimplantation genetic diagnosis to prevent mitochondrial disorders. *Hum Reprod Update* 2008; **14**: 83–94.
2 Bredenoord AL, Dondorp WJ, Pennings G, De Die-Smulders CEM, de Wert G: PGD to reduce reproductive risk: the case of mitochondrial DNA disorders. *Hum Reprod* 2008; **23**: 2392–2401.
3 De Wert GM: The post-menopause: playground for reproductive technology? Some ethical reflections; in Harris J and Holm S (eds): *The Future of Human Reproduction. Ethics, Choice and Regulation*. Oxford: Clarendon Press, 1998, pp 221–237.
4 Pennings G: Measuring the welfare of the child: in search of the appropriate evaluation principle. *Hum Reprod* 1999; **14**: 1146–1150.
5 Bickerstaff H, Flinter F, Yeong CT, Braude P: Clinical application of preimplantation genetic diagnosis. *Hum Fertil* 2001; **4**: 24–30.
6 Steffann J, Frydman N, Gigarel N *et al*: Analysis of mtDNA variant segregation during early human embryonic development: a tool for successful NARP preimplantation diagnosis. *J Med Genet* 2006; **43**: 244–247.
7 Briggs DA, Power NJ, Lamb V, Rutherford AJ, Gosden RG: Amplification of DNA sequences in polar bodies from human oocytes for diagnosis of mitochondrial disease. *Lancet* 2000; **335**: 1520–1521.
8 Dean NL, Battersby BJ, Ao A, Gosden RG, LinTan S, Shoubridge EA: Prospect of preimplantation genetic diagnosis for heritable mitochondrial DNA diseases. *Mol Hum Reprod* 2003; **9**: 631–638.
9 Poulton J, Turnbull DMA: 74th ENMC international workshop: mitochondrial diseases 19–20 November 1999, Naarden, The Netherlands. *Neuromuscul Disord* 2000; **10**: 460–462.
10 White SL, Collins VR, Wolfe R *et al*: Genetic counseling and prenatal diagnosis for the mitochondrial DNA mutations at nucleotide 8993. *Am J Hum Genet* 1999; **65**: 474–482.
11 White SL, Shanske S, McGill JJ *et al*: Mitochondrial DNA mutations at nucleotide 8993 show a lack of tissue or age-related variation. *J Inherit Metab Dis* 1999; **22**: 899–914.
12 Dahl HHM, Thorburn DR, White SL: Towards reliable prenatal diagnosis of mtDNA point mutations: studies of nt8993 mutations in oocytes, fetal tissues, children and adults. *Hum Reprod* 2000; **15** (suppl 2): 246–255.
13 Tatuch Y, Christodoulou J, Feigenbaum A *et al*: Heteroplasmic mtDNA mutation (T-G) at 8993 can cause Leigh disease when the percentage of abnormal mtDNA is high. *Am J Hum Genet* 1992; **50**: 852–858.
14 Ciafaloni E, Santorelli FM, Shanske S *et al*: Maternally inherited Leigh syndrome. *J Pediatr* 1993; **122**: 419–422.
15 Chinnery PF: Inheritance of mitochondrial disorders. *Mitochondrion* 2002; **2**: 149–155.
16 Chinnery PF, Howell N, Lightowlers RN, Turnbull DM: MELAS and MERFF. The relationship between maternal mutation load and the frequency of clinically affected offspring. *Brain* 1998; **121** (Part 10): 1889–1894.
17 Poulton J, Marchington DR: Segregation of mitochondrial DNA (mtDNA) in human oocytes and in animal models of mtDNA disease: clinical implications. *Reproduction* 2002; **123**: 751–755.
18 Blok RB, Gook DA, Thorburn DR, Dahl HHM: Skewed segregation of the mtDNA nt 8993 (T-G) mutation in human oocytes. *Am J Hum Genet* 1997; **60**: 1495–1501.
19 Brown DT, Samuels DC, Michael EM, Turnbull DM, Chinnery PF: Random genetic drift determines the level of mutant mtDNA in human primary oocytes. *Am J Hum Genet* 2001; **68**: 533–536.
20 Mertes H, Pennings G: Oocyte donation for stem cell research. *Hum Reprod* 2006; **22**: 629–634.
21 Pennings G, De Wert GM, Shenfield F, Cohen J, Tarlatzis B, Devroey P: ESHRE task force on ethics and law 12: oocyte donation for non-reproductive purposes. *Hum Reprod* 2007; **22**: 1210–1213.
22 Sandel M: *The Case against Perfection*. Cambridge and London: The Belknap Press of Harvard University Press, 2007.
23 Devolder K: Creating and sacrificing embryos for stem cells. *J Med Ethics* 2005; **31**: 366–370.
24 Dondorp WJ, de Wert GMWR: *Embryonic Stem Cells Without Moral Pain?* The Hague: Centre for Ethics and Health/Health Council of the Netherlands, 2005.
25 Cohen J, Wells D, Munné S: 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* 2007; **87**: 496–503.
26 Combelles CMH: What are the trade-offs between one-cell and two-cell biopsies of preimplantation embryos? *Hum Reprod* 2008; **23**: 493–498.
27 Goossens V, De Rycke M, De Vos A *et al*: Diagnostic efficiency, embryonic development and clinical outcome after the biopsy of one or two blastomeres for preimplantation genetic diagnosis. *Hum Reprod* 2008; **23**: 481–492.
28 Pettman R, Hurley T, Robinson B, Scott H, Kronick JB: Prenatal diagnosis by amniocentesis and chorionic villus biopsy of mtDNA mutation 8993T>G. *J Inherit Metab Dis* 2007; **30**: 404.
29 Chinnery PF, Taylor DJ, Brown DT, Manners D, Styles P, Lori R: Very low levels of the mtDNA A3243G mutation associated with mitochondrial dysfunction *in vivo*. *Ann Neurol* 2000; **47**: 381–384.
30 Jacobs LJAM, De Coo IFM, Nijland JG *et al*: Transmission and prenatal diagnosis of the T9176C mitochondrial DNA mutation. *Mol Hum Reprod* 2005; **11**: 223–228.
31 Harper JC, De Die-Smulders C, Goossens V *et al*: ESHRE PGD consortium data collection VII: cycles from January to December 2004 with pregnancy follow-up to October 2005. *Hum Reprod* 2008; **23**: 741–755.
32 Chen SH, Escudero T, Cekleniak NA, Sable DB, Garrisi MG, Munné S: Patterns of ovarian response to gonadotropin stimulation in female carriers of balanced translocation. *Fertil Steril* 2005; **83**: 1504–1509.
33 Heijnen E, Eijkemans M, De Klerk C *et al*: A mild treatment strategy for *in-vitro* fertilisation: a randomised non-inferiority trial. *Lancet* 2007; **369**: 743–749.
34 Lavery SA, Aurell R, Turner C *et al*: Preimplantation genetic diagnosis: patient's experiences and attitudes. *Hum Reprod* 2002; **17**: 2464–2467.

35 Magli MC, Gianaroli L, Grieco N, Cefalu E, Ruvolo G, Ferraretti AP: Cryopreservation of biopsied embryos at the blastocyst stage. *Hum Reprod* 2006; **21**: 2656–2660.
36 Parriego M, Solé M, Aurell R, Barri PN, Veiga A: Birth after transfer of frozen-thawed vitrified biopsied blastocysts. *J Assist Reprod Genet* 2007; **24**: 147–149.
37 DiMauro S, Hirano M, Schon EA: Approaches to the treatment of mitochondrial diseases. *Muscle Nerve* 2006; **34**: 265–283.
38 Robertson S, Savulescu J: Is there a case in favour of predictive genetic testing in young children? *Bioethics* 2001; **15**: 26–49.
39 Rhodes R: Why test children for adult-onset genetic diseases? *Mt Sinai J Med* 2006; **73**: 609–616.
40 Smith T: *Ethics in Medical Research. A Handbook of Good Practice*. Cambridge: Cambridge University Press, 1999.
41 The World Medical Association: *World Medical Association Declaration of Helsinki. Ethical Principles for Medical Research Involving Human Subjects*. Geneva: World Medical Association, 2008.
42 Council for International Organizations of Medical Sciences: *International Ethical Guidelines for Biomedical Research Involving Human Subjects*. Geneva: CIOMS, 2002.
43 Sue CM, Quigley A, Katsabanis S *et al*: Detection of MELAS A3243G point mutation in muscle, blood and hair follicles. *J Neurol Sci* 1998; **161**: 36–39.
44 Wendler D, Varma S: Minimal risk in pediatric research. *J Pediatr* 2006; **149**: 855–861.
45 Caldwell PHY, Murphy SB, Butow PN, Craig JC: Clinical trials in children. *Lancet* 2004; **364**: 803–811.