

From lab to fertility clinic

Citation for published version (APA):

Jans, V. A. A. M. (2020). *From lab to fertility clinic: the welfare of the child and the ethics of introducing new reproductive technologies*. [Doctoral Thesis, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.20201111vj>

Document status and date:

Published: 01/01/2020

DOI:

[10.26481/dis.20201111vj](https://doi.org/10.26481/dis.20201111vj)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

FROM LAB TO FERTILITY CLINIC

The welfare of the child and the ethics of
introducing new reproductive
technologies

Verna Jans

© Verna Jans, Nijmegen 2020

Printing: ProefschriftMaken II proefschriftmaken.nl

ISBN 978-94-6380-961-0

FROM LAB TO FERTILITY CLINIC

The welfare of the child and the ethics of introducing new reproductive technologies

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit Maastricht,
op gezag van de Rector Magnificus, Prof. dr. Rianne M. Letschert
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op woensdag 11 november 2020 om 12.00 uur

door

Verna Anna Ardina Maria Jans

Promotores

Prof. dr. G.M.W.R. de Wert

Prof. dr. W.J. Dondorp

Copromotor

Prof. dr. H. Mertes (Universiteit Gent)

Beoordelingscommissie

Prof. dr. J.S.M. Krumeich (voorzitter)

Prof. dr. I.D. de Beaufort (Erasmus Universiteit Rotterdam)

Prof. dr. A.L. Bredenoord (Universiteit Utrecht)

dr. R.J.T. van Golde

Prof. dr. T.E. Swierstra

Deze dissertatie doet verslag van een promotieonderzoek dat onderdeel was van het Science and Ethics of stem cell derived Gametes (SEGa) project, gefinancierd door IWT (nu VLAIO), projectnummer 150042. Het onderzoek vond plaats bij de afdeling Health, Ethics and Society (HES) aan de faculteit Health, Medicine en Life sciences (FHML) van Universiteit Maastricht. Het onderzoek was ondergebracht in het programma van GROW, School for Oncology & Developmental Biology.

Table of contents

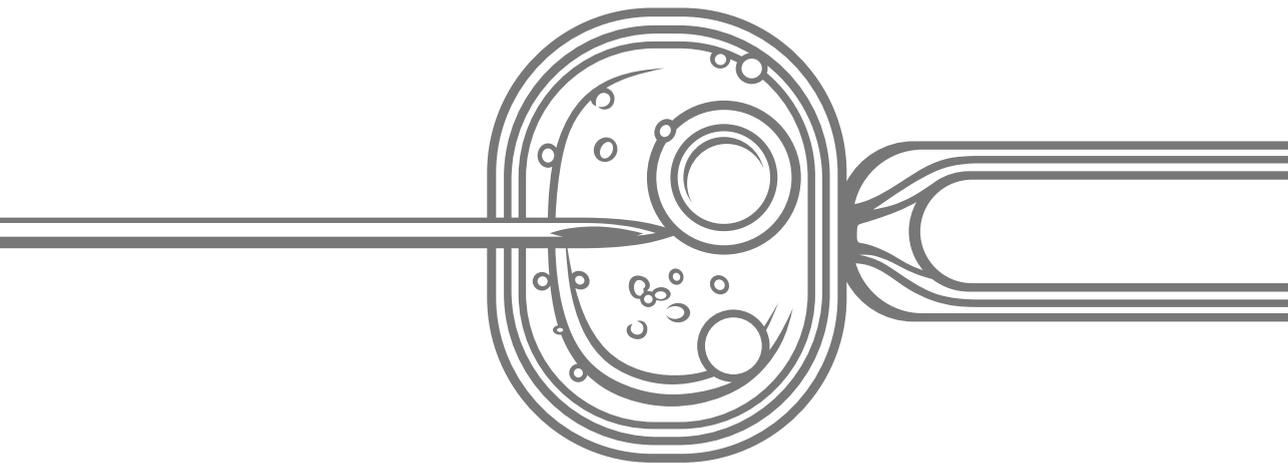
Chapter 1	9
Introduction	
Chapter 2	21
Balancing animal welfare and assisted reproduction: Ethics of preclinical animal research for testing new reproductive technologies	
Chapter 3	43
Of mice and human embryos: Is there an ethically preferred order of preclinical research on new assisted reproductive technologies?	
Chapter 4	59
Between innovation and precaution: How did offspring safety considerations play a role in strategies of introducing new reproductive techniques?	
Chapter 5	87
Follow-up in the field of reproductive medicine: An ethical exploration	
Chapter 6	107
General discussion	

Valorization addendum	141
Samenvatting	145
Acknowledgements	159
About the author	165



Chapter 1

Introduction



1. Background

Assisted reproductive technologies (ARTs) are subject to constant innovation. Since the highly controversial birth of the first 'test-tube baby' about forty years ago, the development of new reproductive techniques increased rapidly and resulted in the dynamic whirlwind the reproductive field is today. Still, while some may say we are already 'playing God', innovation is far from stagnating, as new reproductive technologies swiftly keep popping up on the horizon.

The constant changes in the reproductive field keep instigating a broad range of ethical questions. A first set of questions that arises concerns the moral acceptability of certain reproductive treatments. For example, is the development of a certain reproductive technique at odds with human dignity? Is a certain treatment acceptable when it possibly leads to embryo loss? And what weight should be given to the wish to have a genetic link with one's child? Second, if a certain technique can be considered acceptable in principle, ethical dilemmas present themselves in terms of the conditions to be imposed. For example, to whom should reproductive techniques be offered (what about, for example, singles and applicants affected by a particular addiction)? Also socio-ethical questions are relevant, like should reproductive treatment be reimbursed? Although all of these questions are important to investigate, this thesis will focus on a set of ethical questions concerning how to responsibly deal with medical risks of new reproductive technologies for children thus conceived.

In some fields of medicine, such as the development of drugs, innovation occurs in the context of a strictly regulated system. Here, safety measures are in place in order to protect patients from being exposed to drugs of which the safety and effectiveness is insufficiently proven. In other fields of medicine, however, such as surgical medicine and medically assisted reproduction (MAR), there is no such strictly regulated system. Here, new techniques are often introduced on a basis of trial-and-error, outside a formal research context (Geraedts & de Wert, 2009; Harper et. al., 2012; Margo, 2001). This is especially worrisome in the case of MAR, as 'safety' does not only concern the health and wellbeing of those undergoing the procedures (mostly women), but also that of any children born as a result from these procedures (Pennings et. al., 2007).

Although follow-up research has shown that up until now, only few adverse health effects of new ARTs have emerged, this should not be taken as a reason for complacency, particularly with new reproductive technologies on the horizon that appear to come with potentially significant risks for the offspring. One of such new techniques on the horizon is the reproductive use of stem cell-derived (SCD) gametes (Cyranoski, 2016; K. Hayashi et. al., 2012; K. Hayashi et. al., 2011; Hayashi & Saitou, 2013; Hikabe et. al., 2016; Marques-Mari et. al., 2009; Zhou et. al., 2016).

In the context of this development, the Science and Ethics of SCD Gametes (SEGa) project was founded; a Flemish-Dutch multidisciplinary project aiming at both the scientific development of SCD gametes and timely ethical reflection on both the technique itself and its potential responsible introduction. The first part of this ethical reflection (considering the acceptability of relevant reproductive techniques itself) was performed by Seppe Segers, a member of the ethics team at Ghent University (Segers, 2019). This thesis, concerning the ethics of responsible innovation of MAR, forms the second part of the Ethics Section of this SEGa-project.

There are two potential pathways to create patient-specific SCD gametes for reproductive purposes; 1) by creating an embryo via stem cell nuclear transfer (SCNT) and differentiating the embryonic stem cells (ESCs) into gametes, or 2) by creating induced pluripotent stem cells (iPSCs) and deriving gametes from these (Hendriks et. al., 2015). For both routes, a somatic cell can be used from the patient, which could in turn be differentiated into a gamete and thereafter used to fertilize the gamete of the partner.

If safe and effective, the future reproductive use of SCD gametes would make it possible to better treat several forms of infertility and would thus make donor conception redundant. Using SCD gametes may, for example, enable men with azoospermia, women with primary ovarian insufficiency, people who have had gametotoxic treatments, ageing women with poor quality oocytes and poor responders in IVF programs to produce genetically related offspring.

Additionally, if some serious technical challenges could be overcome, it could give same-sex couples a method to create children of which they are both a biological parent. For example, in a case of an all-male couple, SCD oocytes could be created from pluripotent cells from one of the men. The SCD oocyte could thereafter be fertilized by the naturally conceived

sperm cell from the other man. However, this is more complicated than with heterosexual couples, because the derivation of oocytes from an XY stem line (male) or, even more complicated, sperm from an XX stem line (female), is more likely to have chromosomal problems than derivation of a sex cell from the same sex (de Sutter et. al., 2015).

Although the development of SCD gametes is still at a research stage, research suggests MAR using SCD-gametes is a realistic scenario (particularly for heterosexual couples) in the near to midterm future (Cyranoski, 2016; Hendriks et. al., 2015; Hikabe et. al., 2016; Zhou et. al., 2016). Yet, being a potential risky innovation, irresponsible introduction may lead to harming the health of fertility patients as well as of potentially conceived children by using new ARTs. It may even have transgenerational effects. For example, “the required manipulations and the artificial environments of gametes and embryos are potentially creating short- and long-term health risks in mothers and children by interfering with epigenetic reprogramming” (Grafodatskaya et. al., 2013). This raises the question of whether and how SCD gametes should be responsibly introduced into clinical practice. Moreover, as the potential reproductive use of SCD gametes is only one of many innovations on the horizon, it should be questioned how new ARTs should be responsibly introduced into clinical practice in general.

There is a widely shared agreement that professionals providing assisted reproductive treatments have some level of co-responsibility for the welfare of the future children that they help bring into being (Dondorp & de Wert, 2011; Pennings et. al., 2007). This has led to the European Society of Human Reproduction and Embryology (ESHRE) recommending a normative framework considering responsible innovation in assisted reproduction (Pennings et. al., 2007). This framework consists of three consecutive research steps: (1) preclinical safety studies using animals and/or and human embryos, (2) introducing new techniques in a clinical research setting with registration of outcomes, and (3) long-term follow-up research. It is, however, unclear what this should mean in practice for specific cases and how the successive steps can be made operational taking account of relevant ethical aspects. This thesis aims to fill this lacuna by ethically assessing the current practice of responsible innovation in MAR and by providing practical recommendations concerning how the field of MAR should responsibly deal with safety risks of new reproductive techniques.

2. Research questions and outline of the thesis

The main question of this thesis is: how should the field of MAR responsibly deal with safety risks related to introducing new reproductive techniques? To answer this, several sub-questions are raised alongside the consecutive research steps recommended in the normative framework as proposed by ESHRE.

First, is the use of animals for research into new ARTs acceptable? In the field of MAR, there is a growing emphasis on the importance of introducing new assisted reproductive technologies (ARTs) only after thorough preclinical safety research, including the use of animal models. At the same time, there is international support for the three R's (replace, reduce, refine), and the European Union even aims at the full replacement of animals for research. The apparent tension between these two trends underlines the urgency of an explicit justification for the use of animals for the development and preclinical testing of new ARTs. This question is discussed in chapter 2.

Second, is there an ethically preferred order of preclinical research steps? It is widely acknowledged that the responsible introduction of new assisted reproductive technologies (ARTs) requires preclinical safety research, including the use of animal models and human embryos. However, the moral sensitivity of human embryo research has led to regulations and guidance stating that human embryos may only be used for research that cannot be conducted with animals. In the field of ART research, this translates into the notion of an 'ideal chain' of consecutive preclinical research steps, where research using human embryos may only be considered (when possible) as a further step after promising results have been obtained in animal models first. In chapter 3, I explore the possible moral arguments behind this ideal chain and test their validity.

Third, how did offspring safety considerations play a role in strategies of introducing new ARTs and should this be improved? The field of reproductive medicine has been criticized for introducing ARTs without systematic research on possible safety risks and for failing to meet the standards of evidence-based innovation held elsewhere in medicine. In chapter 4, I firstly ask whether 'responsible innovation' has been a concern for the field, and if so, how it has understood the practical implications of this idea for the development and introduction of potentially risky new ARTs. Secondly, I consider whether the field has lived

up to its responsibilities in this respect, and if not, how things can be improved. To answer these questions, I present three case studies involving the introduction of a new reproductive technology. First, intracytoplasmic sperm injection (ICSI), which is an in vitro fertilization (IVF) procedure in which a single sperm cell is injected directly into the cytoplasm of an egg. Second, preimplantation genetic testing (PGT), which enables the selection of embryos based on an evaluation of their genetic composition. Last, mitochondrial replacement therapy (MRT), which is a form of IVF in which the future child's mitochondrial DNA comes from a donor.

Fourth, what ethical issues arise in follow-up research in the field of MAR? With potentially risky new ARTs on the horizon and with long-term safety data still being scarce even for existing ARTs, follow-up is increasingly seen as essential to responsible innovation in the field of MAR (Harper et. al., 2012; Provoost et. al., 2014; Van Steirteghem, 2008). Despite a general consensus on the importance of follow-up, still very few centers actually conduct it. Both practical and ethical concerns may hamper follow-up being performed responsibly. In chapter 5, I perform an ethical exploration of follow-up research in MAR. Without claiming to be exhaustive, I explore some of the most important ethical implications, issues and concerns related to performing follow-up studies on assisted reproductive technologies (ARTs), taking account of relevant experiences of the academic medical centers of Brussels (Free University), Belgium and Maastricht, the Netherlands.

Last, in the general discussion (chapter 6), I give a summary of all chapters, after which I discuss the findings of this thesis in a wider context. Furthermore, I provide practical recommendations regarding the adjustment and refinement of ESHRE's framework concerning responsible innovation in MAR.

3. Methodology

This thesis is written in the context of applied ethics, specifically biomedical ethics. Both theoretical frameworks and empirical data have been included to establish an 'evidence informed' normative analysis.

Literature research

This thesis aims to perform a normative ethical assessment on how the field of MAR should responsibly deal with safety risks related to introducing new reproductive techniques. Therefore, literature is chosen from scientific and applied ethics journals focusing on technical and normative aspects of reproductive and genetic medicine. Furthermore, books and government documents, from countries both inside and outside of Europe, were used to investigate the international debate and legislation of reproductive medicine.

Empirical research

To collect background information on the introduction of the many reproductive techniques discussed in this thesis, multiple individual interviews were performed with experts in the field of MAR. In the context of MRT specifically, multiple interviews have taken place with experts from the Newcastle Center of Life and Newcastle University, which I was able to visit in November 2018. In 2017, the HFEA awarded the Newcastle group the first UK clinical license to perform MRT, which resulted in the first clinical application in 2018. This group has accepted various women with different indications. However, so far, no pregnancy has been reported. Being one of the leading teams on MRT, the goal of visiting the Newcastle group was to investigate the process of the development and clinical introduction of MRT in the UK.

In addition, two focus group interviews were performed. One of these focus groups involved scientists, clinicians and ethicists working in the field of MAR, and took place at ESHRE's annual meeting in Geneva in 2017. The focus group was aimed at examining their views on responsible innovation in MAR, particularly concerning the notion of an 'ideal order' of preclinical research steps. As these (focus group) interviews solely served the purpose of gathering background information, formal results were not published. The other focus group interview was conducted with MAR-patients. The objective was to investigate their attitude towards potentially risky innovations in MAR. Originally, the aim was to conduct multiple focus group interviews with MAR-patients. However, despite multiple attempts to attract participants in collaboration with patient organizations 'Freya' (The Netherlands) and 'De verdwaalde ooievaar' (Belgium), and in collaboration with the university hospital of Ghent, only a limited number of participants responded. As only one focus group interview could be

performed, with the limited number of eight participants, and due to time constraints, results have not yet been published.

Furthermore, this thesis relied upon the expert knowledge of multiple professionals in the practice of MAR. Information concerning the practice of MAR was drawn from Ellen Goossens (VUB), Hubert Smeets (MUMC+), Maryse Bonduelle (VUB) and Christine de Die (MUMC+). This resulted in several collaborations regarding the development and critical revision of the articles in this thesis.

Normative framework: Wide Reflective Equilibrium

To investigate the ethical issues arising in the responsible innovation of MAR, I make use of *Wide Reflective Equilibrium* (WRE) (Daniels, 1979, 2011; Rawls, 1971). This procedural ethical approach was developed by John Rawls (Rawls, 1971) and was later interpreted by Norman Daniels (Daniels, 1979, 2011). In Rawls' *A theory of Justice* (Rawls, 1971), he proposes that we can only determine what principles of justice we should adopt by broadening our circle of beliefs, on condition that this circle is coherent. This means that WRE does not take an extreme position or give authority to certain moral theories or empirical data, but includes a wide range of moral theories and empirical data in order to reach a normative conclusion. Rawls argues that we should include intuitions, normative frameworks and background theories in the assessment. Examples of such theories, or crucial beliefs, concern the nature of persons, psychology or beliefs about justice (Daniels, 2011). Being a theory of coherence, Rawls emphasizes in his later work that WRE is not an account of absolute truth, but only of justification (Rawls, 1993).

The advantage of using WRE is that it enables evaluating moral judgements in a broader field of moral principles and background theories relevant to responsible innovation in MAR. For example, both the principles of biomedical ethics (Beauchamp & Childress, 2001) and theories on the moral value of animals or human embryos can be used. By including a broad range of considerations, WRE distinguishes itself from more foundational approaches, such as consequentialist ethics and deontology, which only focus on a narrow range of particular cases and moral principles. With 'foundational', I mean that such approaches are based on an account of morality as rooted in what are claimed to be universally valid rules, or in a universally valid understanding of the meaning of moral terms (e.g. categorical

imperative, hedons over dolors, etc.). Additionally, in contrast to WRE, foundational approaches do not take practice-based principles and intuitions into account, which leads to an ethical assessment distancing itself from current practice. Since the ethical debate on responsible innovation in MAR is so closely connected to scientific and medical practice, it is important to take all (practice-based) considerations into account. Therefore, WRE is a better fit to the ethical assessment of this thesis than the rather narrow foundational approaches.

Furthermore, the outcome of the procedure of WRE is ideally fair to all parties, because it is the result of a procedure in which the considerations of all parties have been taken into account (Rawls, 1971, p. 104; 1993, p. 72). Instead of evaluating ethical issues in terms of a foundational framework, WRE examines the consistency of ethical arguments within the overall 'equilibrium' of intuitions and theories, and supplements arguments with general practice-based principles and intuitions. It is important that the considerations of all relevant actors are involved, which is called the *inclusiveness criterion* (Doorn, 2011, p. 76). If the considerations of all parties are included in the ethical assessment, the result of the WRE procedure is fair to all parties. This is called *pure procedural justice*: "what is just is specified by the outcome of the procedure, whatever that is" (Doorn, 2011, p. 70). The WRE is therefore ideal to try to reach moral agreement or at least strong consensus between various parties in ethical debates.

References

- Beauchamp, T. L., & Childress, J. F. (2001). *Principles of biomedical ethics*: Oxford University Press, USA.
- Cyranoski, D. (2016). Mouse eggs made from skin cells in a dish. *Nature News*, 538(7625), 301.
- Daniels, N. (1979). Wide reflective equilibrium and theory acceptance in ethics. *The Journal of Philosophy*, 76(5), 256-282.
- Daniels, N. (2011). Reflective equilibrium 2016(7th of July). Retrieved from The Stanford Encyclopedia of Philosophy website: <http://plato.stanford.edu/entries/reflective-equilibrium/#WidRefEqu>
- Dondorp, W., & de Wert, G. (2011). Innovative reproductive technologies: risks and responsibilities. *Human Reproduction*, 26(7), 1604-1608. doi:10.1093/humrep/der112
- Doorn, N. (2011). *Moral responsibility in R&D networks: A procedural approach to distributing responsibilities*. (Doctoral thesis), TU Delft, Delft.
- Geraedts, J., & De Wert, G. (2009). Preimplantation genetic diagnosis. *Clinical genetics*, 76(4), 315-325.
- Grafodatskaya, D., Cytrynbaum, C., & Weksberg, R. (2013). The health risks of ART. *EMBO reports*, 14(2), 129-135.
- Harper, J., Magli, M. C., Lundin, K., Barratt, C. L., & Brison, D. (2012). When and how should new technology be introduced into the IVF laboratory? *Human Reproduction*, 27(2), 303-313. doi:10.1093/humrep/der414
- Hayashi, K., Ogushi, S., Kurimoto, K., Shimamoto, S., Ohta, H., & Saitou, M. (2012). Offspring from oocytes derived from in vitro primordial germ cell-like cells in mice. *Science*, 338, 971-975.
- Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., & Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell*, 146(4), 519-532. doi:10.1016/j.cell.2011.06.052
- Hayashi, K., & Saitou, M. (2013). Generation of eggs from mouse embryonic stem cells and induced pluripotent stem cells. *Nature protocols*, 8(8), 1513-1524.

- Hendriks, S., Dancet, E. A., van Pelt, A. M., Hamer, G., & Repping, S. (2015). Artificial gametes: a systematic review of biological progress towards clinical application. *Hum Reprod Update*, 21(3), 285-296. doi:10.1093/humupd/dmv001
- Hikabe, O., Hamazaki, N., Nagamatsu, G., Obata, Y., Hirao, Y., Hamada, N., Shimamoto, S., Imamura, T., Nakashima, K., Saitou, M. (2016). Reconstitution in vitro of the entire cycle of the mouse female germ line. *Nature*, 539(7628), 299-303.
- Margo, C. E. (2001). When is surgery research? Towards an operational definition of human research. *Journal of Medical Ethics*, 27(1), 40-43.
- Marques-Mari, A., Lacham-Kaplan, O., Medrano, J., Pellicer, A., & Simón, C. (2009). Differentiation of germ cells and gametes from stem cells. *Human Reproduction Update*, 15(3), 379-390. doi:10.1093/humupd/dmp001
- Pennings, G., de Wert, G., Shenfield, F., Cohen, J., Tarlatzis, B., & Devroey, P. (2007). ESHRE Task Force on Ethics and Law 13: the welfare of the child in medically assisted reproduction. *Human Reproduction*, 22(10), 2585-2588. doi:10.1093/humrep/dem237
- Provoost, V., Tilleman, K., D'Angelo, A., De Sutter, P., de Wert, G., Nelen, W., Pennings, G., Shenfield, F., Dondorp, W. (2014). Beyond the dichotomy: a tool for distinguishing between experimental, innovative and established treatment. *Human Reproduction*, 29(3), 413-417. doi:10.1093/humrep/det463
- Rawls, J. (1971). *A Theory of Justice*. Cambridge, Mass.: Harvard University.
- Rawls, J. (1993). *Political liberalism*: Columbia University Press.
- Segers, S. (2019). *In vitro gametogenesis. An inquiry into the main ethical concerns and why they matter*. (Doctoral thesis), Ghent University, Ghent.
- Van Steirteghem, A. (2008). What next for assisted reproductive technology? A plea for an evidence-based approach. *Human Reproduction*, 23(12), 2615-2616.
- de Sutter, P., Sermon, K., Goossens, E., Pennings, G., Bekaert, S., de Wert, G., & Dondorp, w. (2015). *Science and Ethics of stem cell derived Gametes: IWT application for a SBO project with a primary societal finality*. Universiteit Gent, Universitair Ziekenhuis Gent, Vrij Universiteit Brussel, Universitair Ziekenhuis Brussel, Maastricht University.
- Zhou, Q., Wang, M., Yuan, Y., Wang, X., Fu, R., Wan, H., . . . Zheng, Y. (2016). Complete meiosis from embryonic stem cell-derived germ cells in vitro. *Cell Stem Cell*, 18(3), 330-340.

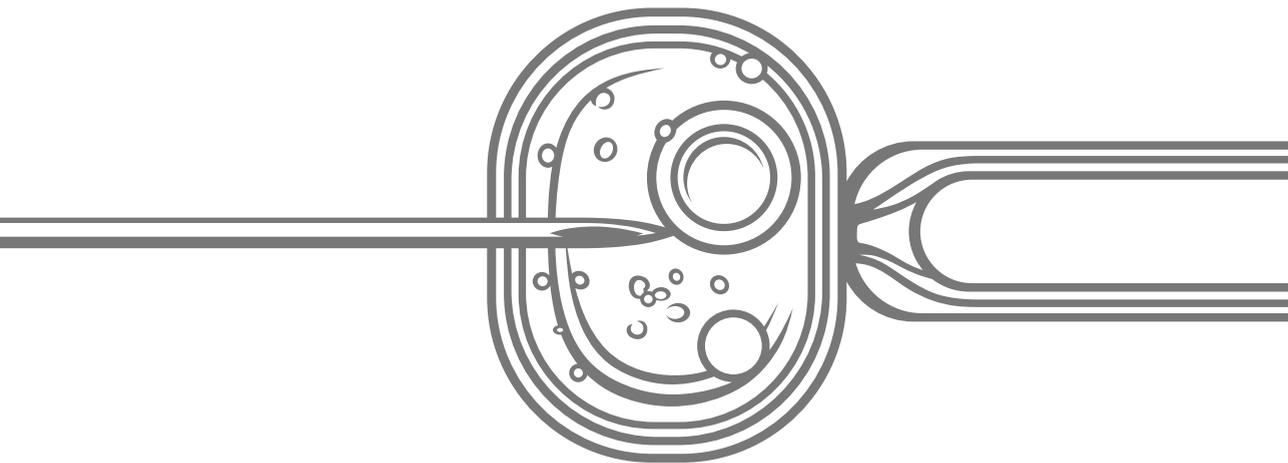


Chapter 2

Balancing animal welfare and assisted reproduction: Ethics of preclinical animal research for testing new reproductive technologies

Published as:

Verna Jans, Wybo Dondorp, Ellen Goossens, Heidi mertes, Guido Pennings, Guido de Wert (2018). Balancing animal welfare and assisted reproduction: ethics of preclinical animal research for testing new reproductive technologies. *Medicine, Health Care and Philosophy*, 21(4), 537-545.



1. Introduction

Several authors have pointed out that all too often, the introduction of new ARTs has taken place on a trial and error basis, thus exposing women and their future children to potential harms of procedures that are possibly risky and of unclear benefit (Dondorp and de Wert, 2011; Harper et. al., 2012; Provoost et. al. 2014; Van Steirteghem, 2008). There is a growing awareness of the importance of making preclinical efficacy and safety research part of the responsible introduction of new assisted reproductive technologies (ARTs). In order to achieve this, the field is called upon to heed existing guidelines (ESHRE) stating that new reproductive technologies should, where possible, be developed and tested in preclinical research using animals and/or human embryos (Pennings et. al., 2007). However, whilst ensuring that new ARTs are safe and beneficial is an ethically recommendable aim, these types of research (on animals and human embryos) are also morally sensitive, thus raising the question how innovation in medically assisted reproduction can responsibly be achieved. In this paper, we will specifically address the ethics of preclinical animal research in this context.

As we will explain in the background section (2), the recent emphasis on the need for such studies in the field of assisted reproduction coincides with simultaneous public concerns and policy efforts in international research governance to minimize the use of animals in scientific research (NC3Rs, 2013; Pijnappel, 2016; Ormandy & Schuppli, 2014; von Roten, 2013). The European Union has even embraced the complete phasing out of animal research as a goal for European legislation (European Commission, 2015). It seems that the call for more rather than less preclinical animal studies in the field of assisted reproduction is out of tune with these policy aims. The least one can say is that this call comes at a time where the justification for the use of animals in research can no longer be taken for granted. The apparent tension between these two trends underlines the urgency of an ethical exploration of whether and under what conditions research involving the preclinical development and testing of new ARTs in animals would be acceptable. Surprisingly, our study is the first to explore this specific question, thus filling an important lacuna relevant to the fields both of assisted reproduction and ethics of animal research.

We start our discussion from the assumption, also underlying the current ethical framework for animal research, that animals may be used for scientifically valid studies under the conditions of subsidiarity and proportionality (section “The ethical framework for animal research”). The first of these principles drives the question whether animals are actually needed to achieve the research aims. We address this issue in section “Alternatives for using animals to test new ARTs”. We argue that given the specific aims of testing new ARTs, a full replacement for using animals is not realistic. The second principle (proportionality) requires an analysis of the moral weight of the aims served by the research held against the moral costs in terms of animal welfare. As will be explained in section “Proportionality of using animals to test new ARTs”, two lines of argument need to be distinguished: the moral weight of helping people to have (genetically own and healthy) children and the importance of making sure that ARTs are only offered when safe and of proven benefit. As we will argue, the strength of the latter argument only counts in the light of a positive assessment of the weight of former. The final section contains concluding remarks.

2. Background

The growing awareness of the importance of preclinical research for new ARTs

As Joyce Harper and colleagues have pointed out, many new ARTs “have not been shown to be safe, to have a clear clinical benefit and/or not been properly validated” before their introduction into clinical practice, which is a “very worrying situation that may become even more common as new technologies are developed” (Harper et. al., 2012, p. 303). In Vitro Fertilization (IVF), Intracytoplasmic Sperm Injection (ICSI), the cryopreservation of embryos, ooplasm transfer, and most recently oocyte vitrification, *in vitro* oocyte maturation and ‘Augment’ treatment are examples of ARTs that have been introduced into clinical practice without much preclinical research (Harper et. al., 2012; Dondorp and de Wert, 2011; Motluk, 2015). However, the past years have seen a growing awareness of the importance of preclinical research on new ARTs.

The general lack of preclinical research has, for example, been recognized by the Task Force on Ethics and Law of the European Society of Human Reproduction and Embryology

(ESHRE). In 2007, they published a statement on the welfare of the child in medically assisted reproduction and pointed out that “the widespread adoption of new techniques frequently takes place without the necessary evaluation of their efficacy (...), safety and social and economic consequences” (Pennings et. al., 2007, p. 13). This exposes women and children to potential risks inherent in the technologies and treatments. These risks include the potential for inducing genetic and epigenetic alterations which may result in lower developmental competence and/or health problems in offspring thus conceived. For example, research has shown that the use of different culture media leads to differences in birthweight in the resulting babies (Dumoulin et. al., 2010).

Animal research in ART

To test the efficacy and safety of new ARTs, ESHRE recommends a research chain consisting of four steps; a) animal studies, b) preclinical embryo research, c) clinical trials and d) follow-up studies (Pennings et. al., 2007). This paper will focus solely on the first step; animal studies. Where animal studies are concerned, ARTs are tested at different stages of reproduction and animal development: studying gametes, preimplantation embryos, fetuses, and offspring in several generations. For these studies, animals are used in different ways: either directly: as the research model to be studied, or indirectly: as providers of those models, when animals are used as a source of gametes or as carriers of fetuses. As animals are needed also for research only looking at gametes or *in vitro* embryos, those studies qualify as ‘animal research’ as well.

The first steps of developing a new ART are often performed *in vitro* (e.g. use of human cell lines or early animal embryos). Thereafter, animal *in vivo* embryos, fetuses and born offspring will be tested. The nature of the precise procedures depends on the research question. For example, in the context of developing Stem Cell Derived (SCD) gametes, animal embryos created with gametes procured with this technology may be used to investigate possible epigenetic effects. After transfer to the womb, animal fetuses grown from such embryos may subsequently be used to test developmental potential. Thereafter, any resulting offspring may be tested for abnormalities, including an evaluation of their reproductive health, and of possible effects upon offspring in further generations.

The insight animal models can give in epigenetic, developmental and transgenerational effects is a significant benefit compared to other types of preclinical research (Brison et. al., 2013). Moreover, in contrast to follow-up research in humans, animal follow-up research can be performed in a relatively short period. In practice, this means that multiple generations of animals generated by a new ART are observed and subjected to several tests during their lives, such as weight measurements and blood tests.

Often, rodents such as mice or rats are used for efficacy and safety research on ARTs, because of their physiological resemblance to humans, the extensive basic knowledge on these animal models and the possibility to perform transgenerational research in a short time span. When the new ART is tested successfully in lower mammals, such as mice or rats, it may be tested on other species (including non-human primates) that allow a better translation of the research outcomes to humans. For example, although Mitochondrial Replacement Therapy (MRT) resulted in the birth of live offspring in the mouse model, when applied on human oocytes, a significant amount of mtDNA from the nuclear donor (patient) was apparently transmitted to the reconstructed embryos. Therefore, the technique would be inappropriate for patients with mtDNA-associated diseases (Tachibana et. al., 2009). To investigate how the transmission of mtDNA could be prevented, researchers have used a non-human primate model, in which they successfully used spindle transfer (ST) whereby no mtDNA was transmitted (Tachibana et. al., 2009).

The Three Rs as a policy aim

Simultaneous with the realization that more preclinical research, including animal research, is needed in the ART field, there is growing international support for the Three Rs (Replace, Reduce and Refine the use of animals in research) (NC3Rs 2013; Pijnappel 2016). Many scientific journals, such as *Human Reproduction*, explicitly request that authors follow the Three Rs (Human Reproduction 2017). According to the principles of the Three Rs, researchers using animals should show that there are no other scientific methods to conduct their research and how they aim to minimize animal numbers and suffering.

In the past few years, the growing support of the Three Rs has resulted in increased regulations on animal research conform to the Three Rs (European Commission, 2010; NC3Rs, 2013). The 2010 European Union's regulations on the protection of animals used for

scientific purposes took full effect in 2013, and replaced the old regulations from 1986. The goal of the new Directive is to “strengthen legislation, and improve the welfare of those animals still needed to be used, as well as to firmly anchor the principle of the Three Rs, to Replace, Reduce and Refine the use of animals, in EU legislation” (European Commission, 2010). The Three Rs are directly incorporated in the Directive, as it states that “animal studies should be either replaced by methods not involving animals, or adapted to reduce the number of animals needed, or refined to minimize pain, suffering or distress experienced by the animal, or to increase their welfare. If an alternative approach to achieve a research objective is available, the Directive makes its use mandatory” (European Commission, 2015). This means that, on a European level, researchers using animals have to show that they incorporate the Three Rs in order to get approval for their research. The ethical justification on this policy will be provided in the next section.

The need for more animal research in the ART field and the growing international support for the Three Rs seem to lead in opposite directions. On the one hand, there is an acknowledged need for doing more animal research as part of the responsible introduction of new ARTs while at the same time the wider message is that we should as much as possible avoid using animals for research. As both these claims refer to ethical principles, the question is how this tension can be resolved in an ethically acceptable way. How should the ART field respond to this tension and what does it mean for safety research using animals on emerging ARTs?

3. The ethical framework for animal research

The acceptability of animal research is a highly debated subject in the field of bioethics as well as in the societal and political domain (Ormandy & Schuppli, 2014; von Roten, 2013). The positions in this debate are often portrayed as being essentially between the biomedical and animal protection communities, diametrically opposing one another (DeGrazia, 2003; Nuffield Council on Bioethics, 2005). Emma Marris, reporter of *Nature*, speaks of “a ping-pong game between hard-core activists and hard-core defenders”, in which “anyone in the middle who stands up to be heard risks getting hit” (Marris, 2006, p. 810). However, in that middle ground, an ethical framework has been developed for the acceptable use of animals in

research for which there is in fact considerable consensus (DeGrazia, 2003; Nuffield Council on Bioethics, 2005).

The ethical framework recognizes that as sentient beings, animals have a moral status that needs to be taken seriously. A being has moral status, if it has interests which can be harmed (Singer, 1995). The moral status of animals means that we are obliged to consider their needs and interests when making decisions affecting them (Warren, 1997). Applying this to the practice of animal research, the framework insists that animals are as much as reasonably possible treated with respect to their well-being, which also includes allowing them to live the typical life of a representative of their species. As the qualifier ('as much as reasonably possible') indicates, not every infringement on animal wellbeing is as such regarded as reprehensible. However, recognition of their moral status entails that any such infringement requires a justification. The obvious candidate justification for animal research is that it would further the interests of humans in the development of new knowledge and better medical technologies. The framework specifies the conditions for this justification in terms of two requirements; 1) that the benefits to be derived from animal research cannot be obtained through other, less morally problematic, means (principle of subsidiarity) and 2) that the level of infringement on animal wellbeing is proportional to the moral weight of those benefits (principle of proportionality) (European Commission, 2010). In practice, this means that an animal ethics committee assessing a protocol for animal research must agree that the research meets both these requirements.

Implicit in this ethical framework is the notion that animal research always comes at a moral price. Its justification therefore requires not only that, with regard to specific research protocols, this price can be shown to be the lesser of two evils (with refraining from research as the larger one), but also that accepting the practice of animal research should come with a general and continuing commitment on the part of society to as much as possible reduce (and ideally avoid) this implicit moral price. This is reflected in the concept of the 3Rs (NC3Rs, 2016). Although the concept, with its emphasis on minimizing the use of animals and improving the welfare of those still being used, was already developed over fifty years ago, it is increasingly seen today as a context for overcoming the moral disagreement between the biomedical and the animal protection communities. (Pijnappel, 2016).

4. Alternatives for using animals to test new ARTs

The principle of subsidiarity drives the question whether animals are actually needed to achieve the research aims. As discussed in the “Background” section, animal research can give insight in the risks of a certain technology and may help to reduce these risks before applying the technology to humans. Are there alternative methods with less invasive means than animal research to investigate new ARTs?

Elsewhere in medicine, alternatives for the use of animals in preclinical research have been found in computer models, the use of alternative organisms, such as lower microorganisms (e.g. yeasts), and the use of animal/human *in vitro* cell and tissue cultures (Doke and Dhawale 2015). Bearing in mind the nature of ART research described in the “Background” section, it is implausible that these models can replace all ART animal research, as they cannot provide information on the development of embryos and fetuses, nor on the health of (multigenerational) offspring. For example, in studies investigating the safety of Spermatogonial Stem Cell Transplantation (SSCT) and Testicular Tissue Grafting (TTG), there were no alternatives for animal research when testing for epigenetic effects in multigenerational offspring (Goossens et. al., 2011; Goossens et. al., 2009). The use of animal/human *in vitro* cell and tissue cultures or the use of e.g. organoids may be an interesting option to reduce the use of animals for research (Lancaster & Knoblich, 2014). However, since many research procedures, such as performing transgenerational research, are impossible in this model, they will only have limited value.

A possibly important method to minimize the use of animals, which has not been given much attention yet as an alternative to animal research in the literature, can be research on human embryos *in vitro* (i.e. preclinical human embryo research). Embryo research enables researchers to test offspring for abnormalities in an early stage of life. For instance, molecular studies on preimplantation embryos can be useful to yield information on epigenetic reprogramming. Moreover, the use of human embryos can also be helpful to avoid translation problems: “Preferentially and where possible, studies should be done on human embryos, as results from animal studies cannot always be extrapolated to humans” (de Rycke et. al., 2002).

Proposing human embryo research as an alternative presupposes that the use of human embryos as research material is morally less problematic than that of animals, since the subsidiarity principle should be met. This, however, seems to run against the widely held view that the opposite is the case: that human embryos have a higher moral status and deserve more respect than animals. For instance, in many jurisdictions that allow human embryo research under conditions, one of those conditions is that the envisaged results cannot be obtained through other methods, including animal research. The British Warnock report explicitly states that human embryos should only be used when research with animal models is not an option (Warnock, 1985). The same view is reflected in ESHRE's account of a research chain in which human embryo studies are described as the next step, only after preceding animal studies have led to reassuring data (Pennings et. al., 2007). Interestingly, the subsidiarity principle underlying the call for replacing animal research also applies to human embryo research. This means that, when holding on to the view that human embryos have a higher moral status than animals, the proposition of using human embryo research to replace animal research collides with the condition to only perform human embryo research when animal research is not an option.

We however do not find it obvious that the use of animal models is indeed morally less problematic than human embryo research. Firstly, animal research inevitably entails an infringement of animal wellbeing whereas human embryos cannot experience pain or discomfort. Secondly, as reflected in the acceptance of occasioning left-over embryos in IVF-practice, the moral status of human preimplantation embryos is relatively low (Harman, 1999; Steinbock, 2011; Gezondheidsraad, 1998). How the conflict between these readings of the subsidiarity principle should be resolved is a matter for further debate and analysis that should also include the possible scenario of extending human embryo research beyond the current 14-day limit (Hyun et. al., 2016), if that could yield important information about the safety of a new technology. To the extent that human embryo research can indeed be considered as morally less rather than more problematic than using animals for the same purpose, present accounts of the order of the ideal chain of types of preclinical safety research need to be revised. Even so, as human embryo research will not provide data beyond the earliest stages of development, it cannot be expected to provide more than a partial replacement of the use of animals for testing new ARTs.

5. Proportionality of using animals to test new ARTs

The conclusion that animal research remains necessary for the responsible introduction of new ARTs is only a first step in determining the ethical acceptability of such research. The next step concerns the question whether the aims served with ART research are important enough to justify the use of animals. Whereas in the previous section we explored whether the subsidiarity principle was met, we now turn to the question of whether such research is also proportional. To answer this question, we need to explore what is in the scales on both sides of the balance. On the moral cost side, we first need to get a clearer picture of what precisely is at stake in terms of the impact on animal wellbeing where testing animals for ART research is concerned. We will then move on to the other side of the balance in order to discuss whether the moral benefits of ART research are substantial enough to outweigh the moral cost.

The moral cost of using animals for ART research

In the “Background” section, we described the research process using animals in different stages of their lives. How do these research procedures translate to moral cost in terms of compromised animal welfare and animal suffering? To perform animal gamete or embryo research *in vitro*, oocytes and sperm need to be isolated from full grown animals. In some species, such as cattle, researchers can use spare material from abattoirs to retrieve gametes. In most other species, however, such as in the mouse model, females must undergo hormone stimulation. The injections lead to short and light pain for the animals, but the hormone stimulation itself does not lead to animal distress. The mice will be euthanized and their reproductive tissues removed, after which oocytes are collected from the fallopian tubes of females and sperm from the epididymis of males. Hormonal stimulation of the females is also a first step in research involving animal embryos *in vivo*. A few days after mating, the carrier will be euthanized and its fallopian tubes and uterus removed in order to allow the isolation of fertilized oocytes (or embryos). Fetuses and newborn animals will be euthanized prior to being tested, for instance for epigenetic abnormalities. To test the adults, a piece of tissue will be removed which may be slightly painful, but that is something that can be avoided by sedation. Thereafter, the tissue will be tested. Ultimately, most research animals will be

euthanized when they are no longer useful for the research (except for e.g. nonhuman primates).

How do we evaluate the moral cost of these procedures? If one wants to assess animal research for a specific research project, the moral cost depends on what procedures are needed for answering the precise research question and concerns elements such as the number of animals used, the use of anesthetics and which animal species are needed. Here, we want to assess the moral cost of animal research on ARTs in comparison to other practices using animals. Several accepted animal research practices are highly invasive. For instance, in studies aimed at developing therapies for cancer or burn injuries, animals are exposed to conditions that may cause significant suffering as a result of tumor growth or inflicted burns (Abdullahi et. al., 2014; Workman et. al., 2010). Animal research on ARTs, on the contrary, is much less invasive, since animals will be at most subjected to short and light pain and are killed painlessly instead of being subjected to significant suffering. As explained by philosopher Jeff McMahan, the general view is that “it is more important to prevent the significant suffering of animals than it is to prevent, or not to cause, their deaths” (McMahan, 2002). We conclude that the harm done to animals in ART research is, in comparison to cancer or burn research, relatively low, resulting in a relatively low moral cost.

Comparing animal research on ARTs to another widely accepted practice of animal use, namely the social practice of eating meat, invites the same conclusion that animal research on ART has a relatively low moral cost. Whereas animal research should always meet strict rules in accordance with the principles of the Three Rs, these principles are not incorporated by the food industry. Consequently, research animals are generally treated with more respect and are subjected to less suffering (e.g. using anesthetics, proper housing) than food industry animals.

As animal research on ARTs comes with a relatively low level of harm, the moral cost of these procedures is low in comparison with that of other generally accepted animal research practices and with the use of animals in the meat industry. Nevertheless, since also in animal research on ARTs, animals suffer some degree of pain or are killed, their use as research models is in need of justification. This leads to the question whether the moral cost is proportional to the potential benefits.

Justifying aims of animal research: human health as a benchmark

In line with the ethical framework discussed in section "The ethical framework for animal research, the European Directive refers to "the avoidance, prevention, diagnosis or treatment of disease, ill-health or other abnormality or their effects in human being" as justifying aims of animal research (European Commission, 2010). This reflects the widely shared notion that health is such a vital need for humans that if animal research can ever be regarded as serving a sufficiently important aim, it must be for health. If we take this as the upper end of a spectrum of considered acceptability, then most cosmetics research would be found at the lower end. Longstanding ethical concerns about the proportionality of the use of animals for consumer cosmetics has led to an official ban of this practice in the entire European Union since 2013 (European Commission, 2009). Apart from the availability of possible alternative methods for safety-testing, an important reason for this is that developing products that merely cater to consumer preferences is not regarded as sufficiently weighty to justify the inevitable infringement of animal wellbeing. Taking the uncontested value of health as a benchmark, our question thus becomes where on this spectrum the benefits provided by ART research are to be positioned.

This question cannot simply be answered by pointing to the fact that the relevant research is meant to protect humans from possible health risks connected to untested technologies. While this is a necessary element of any justification of the use of animals for preclinical safety studies, it is not sufficient to make the case. As with regard to consumer cosmetics, whether the use of the relevant technology would be safe only becomes a proportionality affecting issue when it is first established that the use of the technology as such serves a morally weighty aim. Therefore, the proportionality question should be addressed on two levels. We will start our assessment on the first level, where it is questioned whether the benefits of MAR are important enough to justify the use of animals. Only when this criterion can be met, the step can be made to the second level, where it is questioned whether the aims of ART *research* are important enough to justify the use of animals.

The moral importance of medically assisted reproduction

With regard to the first level, the aim for which MAR was originally developed, i.e. helping people with fertility problems to have children, has led to a long debate about whether

infertility should be regarded as a disease (Holm 1996). Reasons for an affirmative answer (Zegers-Hochschild et. al., 2017) are that on the level of biological functioning, fertility problems can be ascribed to observable or presumed abnormalities in the reproductive system. Although MAR does not take away or 'cure' those abnormalities, it provides fertility patients with something (a child) for which reproductively healthy couples need no medical help. Clearly, for those taking this view, developing new ARTs for people with fertility problems is a matter of developing health care, putting the importance of the needed research in the higher end of the spectrum.

Others have however argued that infertility is not so much a disease, but rather a (social) handicap (Holm, 1996). In their view, childlessness becomes a problem in the context of societal expectations and personal desires, rather than as a direct consequence of a biological abnormality. This also connects with the idea that for those who regard their childlessness as a problem, there are other options for having a child, thus relativizing the importance of developing new ARTs. For example, adoption is an alternative to medical treatment, or having a child through low-tech donor insemination is an alternative to high-tech ICSI. This might lead to the conclusion that the importance of developing new ARTs is relatively low. However, the value that many people place on having a child with a genetic link to (ideally) both partners, points in the opposite direction and is an important driver for developing new ARTs capable of providing just that: a child of which both partners are the genetic parents.

Currently, ARTs are not only being developed to help the infertile to have genetically related children, but also to help other people, regardless of their fertility status. Future MAR options using SCD-gametes will to a large extent serve that aim. This might also make it possible for people who are unable to reproduce due to their sexual orientation, relationship status or age to have a genetically related child. For women who for whatever reason expect not to be able to fulfill their child-wish prior to running out of functional oocytes, the development of SCD-gametes will also make oocyte or ovarian tissue cryopreservation redundant. Since it is obvious that when dealing with this wider range of requests, MAR does not respond to a health problem and that for those involved there may be alternative routes to having a child, the question becomes how the importance of these reproductive services

relates to the uncontested value of human health on the one end of the spectrum and the trivial importance of consumer cosmetics on the other.

A specific form of MAR entails helping people to have children who are not only genetically related, but also healthy. This concerns people who may or may not have a fertility problem, but who want to avoid having a child with a genetic disorder that they are at risk of transmitting. MAR treatments that will allow them to have a child without the specific genetic disease include PGD, MRT, or possibly germline gene editing (CRISPR-cas9) in the future. In terms of our spectrum, treatments to avoid the transmission of genetic diseases might be regarded as falling in the category of health. The argument behind this reasoning is that a healthy child can only be realized by ensuring that the child is born without the genetic defects of which their parents are healthy or affected carriers. However, to realize the aim of having a healthy child, parents can also choose for using donor gametes or adoption instead of medical treatment. Nevertheless, many patients choose the latter option, despite the invasiveness of ART treatments. This shows, again, that there is apparently a socially determined motive for preferring the options leading to a genetically related child over other alternatives.

As it appears most appropriate to say that the practice of MAR serves the aim of helping people to have a genetically related child, our question about the moral weight of developing new ARTs requires an assessment of the importance of this aim. Although research has shown that there are no significant differences considering psychological well-being between biologically and non-biologically related parents and children (Golombok et. al., 2006; Lansford et. al., 2001; Golombok et. al., 2004), and although some commentators have called the preference to have genetically related children ‘irrational’ (Holm, 1996; Bayles, 1984), it is clear that many people do consider the genetic link to be very important. For instance, the decision to establish a family through donor conception is often only made after a long process of failed ART treatments with the prospective parents’ own gametes and coming to terms with this situation may for many be possible only after a period of serious grieving. Moreover, what are often presented as alternative ways of having a child are not always available due to strict conditions (adoption) or scarcity (donor oocytes) or come with challenges of their own. For example, the abolishment of donor anonymity in a growing number of countries (which in itself can be regarded as reflecting an increased societal

emphasis on the importance of the genetic link (Pennings, 2012)) has made donor conception less attractive to many couples who are weary about future contact of their child with the donor and the possible role that the donor may want to play in their family (Brewaeys et. al., 2005).

We acknowledge that given the different views about the importance of the genetic link, MAR cannot be said to serve a vital human need in the same way some other medical interventions do. However, it would neither be correct to say that it serves a trivial preference. Given that those different views connect to the plurality of understandings of what a flourishing life means, and in the light of the value attached to that plurality in our liberal society, we argue that the practice of MAR does indeed serve a morally weighty aim.

Is MAR important enough to justify the use of animals in research aimed at improving the practice (if no alternative methods are available)? We see two reasons for arguing that it is. Firstly, as long as the large-scale use of animals for food, where alternatives are readily available in the form of vegetarian or vegan lifestyles, is considered morally acceptable by society, it seems difficult to maintain that research aimed at improving MAR would not be sufficiently important. Secondly, the earlier observation that, also in comparison with the use of animals in the meat industry, preclinical ART research has a relatively low impact on animal wellbeing, further supports this conclusion.

The proportionality of preclinically testing new ARTs

If the above reasoning is sound, the argument on the next level can be more straightforward. Here we are concerned with the moral importance of preclinically testing new ARTs for efficacy and safety. Testing the efficacy of a new ART prior to its clinical introduction is important to avoid subjecting women (and sometimes men) to potentially risky treatment without sufficient benefit. Testing its safety also concerns the health of children and future generations. Although animal research can never guarantee the safety of a new technology, testing ARTs in animal models prior to introducing them in the clinic can reduce the risk that children conceived through these technologies may be born with malformations or experience health problems during the rest of their lives. The fact that only few such health effects of new ARTs have until now emerged cannot be a reason for complacency in this regard (Dondorp & de Wert, 2011). As embryologist Anne McLaren has famously said, the

direct clinical introduction of untested new ARTs can be compared with ‘making the first test of a new aircraft-guidance system on a crowded Boeing 747’ (McLaren, 1989). As this regards human health related concerns that as such fall on the higher end of the spectrum of justificatory aims of animal research, we conclude that a strong case can be made that, in principle, the use of animals for testing new ARTs would be proportional.

6. Concluding remarks

Considering that the use of animals remains necessary for specific forms of ART research and taking account of different views of the moral importance of helping people to have a genetically related child, we argue that, in principle, the importance of safety research as part of responsible innovation outweighs the limited infringement of animal wellbeing involved in ART research. In principle, because the proportionality of concrete instances of animal research still depends on the specifics of the case. As animal research inevitably comes at a moral price, it remains important to constantly consider to what extent the ideal of the 3Rs can be better met. Although a full replacement of animal research on new ARTs is not realistic, possible alternatives including the use of human embryos should be considered in concrete cases.

References

- Abdullahi, A., Amini-Nik, S., & Jeschke, M.G. (2014). Animal models in burn research. *Cellular and molecular life sciences* 71 (17):3241-3255.
- Bavister, B. D., Wolf, D. P., & Brenner, C. A. (2005). Challenges of primate embryonic stem cell research. *Cloning and stem cells* 7 (2):82-94.
- Bayles, M. D. (1984). *Reproductive ethics*. Upper Saddle River, NJ: Prentice-Hall.
- Brewaeys, A., de Bruyn, J. K., Louwe, L. A., & Helmerhorst, F. M. (2005). Anonymous or identity-registered sperm donors? A study of Dutch recipients' choices. *Human Reproduction* 20 (3):820-824.
- Brisson, D. R., Roberts, S. A., & Kimber, S. J. (2013). How should we assess the safety of IVF technologies? *Reproductive Biomedicine Online* 27 (6):710-721. doi:10.1016/j.rbmo.2013.09.006.
- European Commission. (2009). Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetics. *Official Journal of the European Union product L* 342: 52-209.
- Cobo, A., Meseguer, M., José Remohí, & Pellicer, A. (2010). Use of cryo-banked oocytes in an ovum donation programme: a prospective, randomized, controlled, clinical trial. *Human Reproduction* 25 (9):2239-2246.
- Council of Europe. (2005). European convention for the protection of vertebrate animals used for experimental and other scientific purposes, ed. Council of Europe.
- DeGrazia, D. (2003). The Ethics of Animal Research: What are the Prospects for Agreement? In *The Animal Ethics Reader*, eds. Susan J. Armstrong, and Richard G. Botzler, 309-316. Abingdon: Routledge.
- Doke, S. K., & Dhawale, S. C. (2015). Alternatives to animal testing: A review. *Saudi Pharmaceutical Journal* 23 (3):223-229.
- Dondorp, W. & de Wert, G. (2011). Innovative reproductive technologies: risks and responsibilities. *Human Reproduction* 26 (7):1604-1608. doi:10.1093/humrep/der112.
- Dumoulin, J. C., Land, J. A., van Montfoort, A. P., Nelissen, E. C., Coonen, E., Derhaag, J. G., Schreurs, I. L., Dunselman, G. A., Kester, A. D., Geraeds, J. P. & Evers, J. L. (2010). Effect

- of in vitro culture of human embryos on birthweight of newborns. *Human reproduction* 25 (3):605-612.
- European Commission. (2010). Directive 2010/63/EU on the protection of animals used for scientific purposes. In *2010/63/EU*, ed. European Commission. Ispra, Italy.
- European Commission. (2015). Communication from the commission on the European Citizens' Initiative "Stop Vivisection". Brussels.
- Gezondheidsraad. (1998). IVF: afrondende advisering Rijsewijk: Gezondheidsraad: Commissie Herziening Planningsbesluit IVF.
- Golombok, S., Lycett, E., MacCallum, F., Jadvá, V., Murray, C., Rust, J., Abdalla, H., Jenkins, J., & Margara, R. (2004). Parenting infants conceived by gamete donation. *Journal of Family Psychology* 18 (3):443.
- Golombok, S., Murray, C., Jadvá, V., Lycett, E., MacCallum, F., & Rust, J. (2006). Non-genetic and non-gestational parenthood: consequences for parent-child relationships and the psychological well-being of mothers, fathers and children at age 3. *Human Reproduction* 21 (7):1918-1924.
- Goossens, E., Bilgec, T., van Saen, D., & Tournaye, H. (2011). Mouse germ cells go through typical epigenetic modifications after intratesticular tissue grafting. *Human Reproduction* 26 (12):3388-3400.
- Goossens, E., de Rycke, M., Haentjens, P., & Tournaye, H. (2009). DNA methylation patterns of spermatozoa and two generations of offspring obtained after murine spermatogonial stem cell transplantation. *Human Reproduction* 24 (9):2255-2263.
- Harman, E. (1999). Creation ethics: The moral status of early fetuses and the ethics of abortion. *Philosophy & public affairs* 28 (4):310-324.
- Harper, J., Magli, M. C., Lundin, K., Barratt, C. L., & Brison, D. (2012). When and how should new technology be introduced into the IVF laboratory? *Human Reproduction* 27 (2):303-313. doi:10.1093/humrep/der414.
- Hewitson, L. (2004). Primate models for assisted reproductive technologies. *Reproduction* 128 (3):293-299.
- Holm, S. (1996). Infertility, childlessness, and the need for treatment. *Creating the Child*. Dordrecht: Kluwer:65-78.

- Human Reproduction. (2018). Journals policies and precedures. <https://academic.oup.com/humrep/pages/Policies>. Accessed January 23rd 2018
- Hyun, I., Wilkerson, A. & Johnston, J. (2016). Embryology policy: Revisit the 14-day rule. *Nature* 533:169-171.
- Lancaster, M. A. & Knoblich, J. A. (2014). Organogenesis in a dish: modeling development and disease using organoid technologies. *Science* 345 (6194):1247-125.
- Lansford, J. E., Ceballo, R., Abbey, A., & Stewart, A. J. (2001). Does family structure matter? A comparison of adoptive, two-parent biological, single-mother, stepfather, and stepmother households. *Journal of Marriage and family* 63 (3):840-851.
- Liebsch, M., Grune, B. Seiler, A., Butzke, D., Oelgeschläger, M., Pirow, t., Adler, S., Riebeling, S., & Luch, A. 2011. Alternatives to animal testing: current status and future perspectives. *Archives of Toxicology* 85 (8): 841-858.
- Marris, E. (2006). Grey Matters. *Nature* 444 (7121): 808-810
- McLaren, A. (1989). IVF: regulation or prohibition? *Nature* 342 (6249):469-470.
- McMahan, J. (2002). *The ethics of killing: Problems at the margins of life*. Oxford: Oxford University Press.
- Motluk, A. (2015). IVF booster offered in Canada but not US. *Canadian Medical Association Journal* 187 (3):E89-90. doi:10.1503/cmaj.109-4975.
- NC3Rs. (2013). 2013 Research Review: National Centre for the Replacement, Refinement and Reductions of Animals in Research.
- NC3Rs. (2016). The 3Rs. <https://www.nc3rs.org.uk/the-3rs>. Accessed November 10th 2016.
- Nuffield Council on Bioethics. (2005). The ethics of research involving animals.
- Ormandy, Elisabeth H, and Catherine A Schuppli. (2014). Public attitudes toward animal research: a review. *Animals* 4 (3):391-408.
- Pennings, G., de Wert, G., Shenfield, F., Cohen, J., Tarlatzis, B., & Devroey, P. (2007). ESHRE Task Force on Ethics and Law 13: the welfare of the child in medically assisted reproduction. *Human Reproduction* 22 (10):2585-2588. doi:10.1093/humrep/dem237.
- Pennings, G. (2012). How to kill gamete donation: retrospective legislation and donor anonymity. *Human Reproduction* 27 (10):2881-2885.

- Pijnappel, M. C. (2016). *Lost in technification: Uncovering the latent clash of societal values in Dutch public policy discourse on animal-testing alternatives*. Thesis Radboud University Nijmegen. Enschede: Ipskamp Drukkers. <http://repository.ubn.ru.nl/handle/2066/151524>. Accessed June 5th 2017.
- Provoost, V., Tilleman, K., D'Angelo, A., de Sutter, P., de Wert, G., Nelen, W., Pennings, G., Shenfield F., & Dondorp, W. (2014). Beyond the dichotomy: a tool for distinguishing between experimental, innovative and established treatment. *Human Reproduction* 29 (3):413-417. doi:10.1093/humrep/det463.
- von Roten, F. C. (2013). Public perceptions of animal experimentation across Europe. *Public Understanding of Science* 22 (6):691-703.
- de Rycke, M, Liebaers, I, & van Steirteghem, A. (2002). Epigenetic risks related to assisted reproductive technologies: Risk analysis and epigenetic inheritance. *Human Reproduction* 17 (10):2487-2494,
- Singer, P. (1995). *Animal liberation*. New York, NY: Random House.
- Steinbock, B. (2011). *Life before birth: the moral and legal status of embryos and fetuses*. Oxford: Oxford University Press.
- van Steirteghem, A.. (2008). What next for assisted reproductive technology? A plea for an evidence-based approach. *Human Reproduction* 23 (12):2615-2616.
- Tachibana, M., Sparman, M., Sritanandomchai, H., Ma, H., Clepper, L, Woodward, J., Li, Y., Ramsey, C., Kolotushkina, O., & Mitalipov, S. (2009) Mitochondrial gene replacement in primate offspring and embryonic stem cells. *Nature* 461 (7262): 367-372
- Warnock, M. (ed). (1985). *A question of life: the Warnock report on human fertilisation and embryology*. Hoboken, NJ: Wiley-Blackwell
- Warren, M. A. (1997). *Moral status: Obligations to persons and other living things*. Oxford: Clarendon Press.
- Workman, P., Aboagye, E. O., Balkwill, F., Balmain, A., Bruder, G., Chaplin, D. J., Double, J. A., Everitt, J., Farningham, D. A. H., & Glennie, M. J. (2010). Guidelines for the welfare and use of animals in cancer research. *British journal of cancer* 102 (11):1555-1577.
- Zegers-Hochschild, F., Adamson, g. D., Dyer, S., Racowsky, C., de Mouzon, J., Sokol, R., Rienzi, L., Sunde, A., Schmidt, L., & Cooke, I. D. (2017). The International Glossary on Infertility and Fertility Care, 201. *Human Reproduction* 32 (9): 1786-1801.

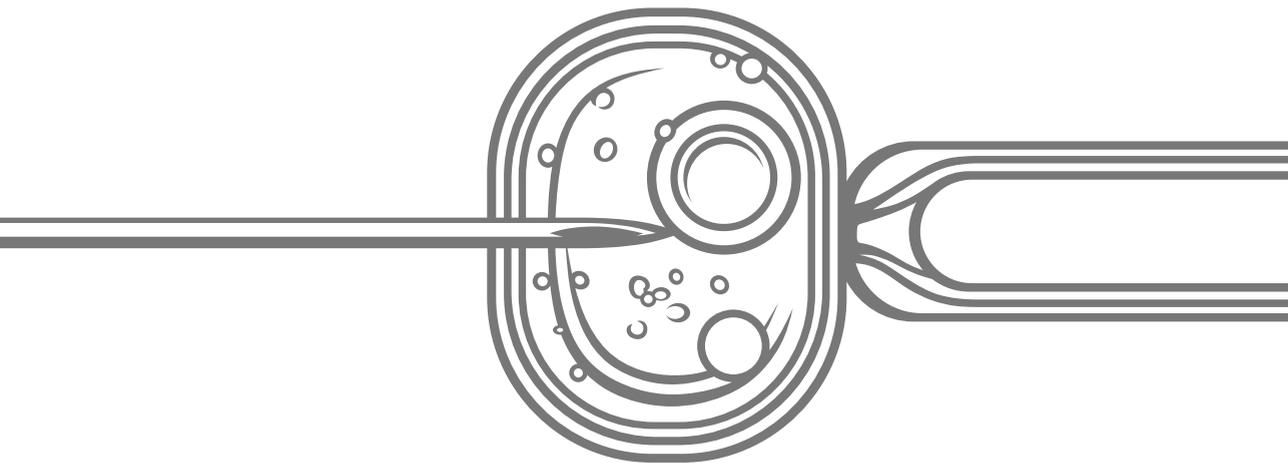


Chapter 3

Of mice and human embryos:
Is there an ethically preferred order of
preclinical research on new assisted
reproductive technologies?

Published as:

Verna Jans, Wybo Dondorp, Ellen Goossens, Heidi Mertes, Guido Pennings, Hubert Smeets, Guido de Wert (2018). Of mice and human embryos: is there an ethically preferred order of preclinical research on new assisted reproductive technologies? *Human Reproduction*, 33(9), 1581-1585.



1. Introduction

An important concern with developments in medically assisted reproduction is that new techniques or add-ons to existing ones are often introduced without sufficient preclinical safety and efficacy studies, and also without meaningful efforts to collect follow-up data. As several authors have pointed out, the field fails to meet standards of evidence-based innovation held elsewhere in medicine (Dondorp & De Wert, 2011; Harper et. al., 2012; Harper et. al. 2017; Sharpe, 2018). This is problematic not just in terms of the good-quality medical care that fertility patients may expect for themselves, but also because possible risks of new assisted reproductive technologies (ARTs) extend to the offspring conceived with these methods. In the light of this, one would expect the field to take the safety of its procedures far more seriously than it currently does. The point was forcefully brought home in a recent review in this journal of the emerging awareness of the relevance of the ‘developmental origins hypothesis of health and disease’ for responsible innovation in reproductive medicine. As the author observes, “it seems clear that the techniques applied in ART do actually affect the growth, development and health of offspring, but it is unclear to what extent each of the techniques does so, what the effects are on short and long-term health and what the optimal treatment would be” (Roseboom, 2018).

In the light of these concerns, there is growing support for the view that new reproductive technologies should only be introduced after predictive preclinical efficacy and safety studies (Harper et. al., 2012; Mulder et. al., 2018; Pennings et. al., 2007; Provoost et. al., 2014; Sharpe, 2018; van Steirteghem, 2008). For preclinical research in this field, cellular models, animal models (ranging from rodents to non-human primates) and human embryos can be used, depending on the research question. However, the moral sensitivity of human embryo research has led to regulations and guidance requiring that the use of human embryos may only be considered for important (basic or preclinical) research questions that cannot be answered through other means, including animal models. We call this the ‘use animals first’ (UAF) rule.

At the same time, animal research is morally sensitive as well, as is reflected in the growing support of the so-called 3Rs, calling researchers to ‘replace, reduce, refine’ the use of research animals wherever possible (NC3Rs, 2018). Interestingly, the ethical framework

for animal research insists that animals may only be used for research questions that cannot be answered through other means (Jans *et. al.*, 2018). As in this context ‘other means’ would seem to include research using human embryos, researchers may find themselves in a *catch-22* situation regarding the prioritization of these two types of preclinical research. In this paper, we explore the possible moral arguments behind the UAF-rule and test their validity. Although this rule applies to multiple scientific practices, the scope of this paper is limited to preclinical ART research. When referring to human embryos as possible research material, we mean early (<14 days) in vitro embryos of sufficient quality that are either donated to research during an IVF-treatment procedure, or embryos especially created for research (‘research embryos’).

2. The ‘use animals first’ rule

As a general rule, the UAF-rule takes the form of a subsidiarity requirement that applies to all forms of human embryo research, regardless of context and purpose. An early example of this can be found in the seminal Warnock report (Warnock, 1985) that has not only stood at the cradle of the British legislation of human embryo research, but has influenced relevant regulations in other countries as well. In this report it is said that “no one should undertake research on human embryos the purposes of which could be achieved by the use of other animals or in some other way” (p. 63). The argument given for upholding this rule is that “the human embryo is entitled to some added measure of respect beyond that accorded to other animal subjects” (p. 62). A similar position was taken by the Health Council of the Netherlands in an advisory report commissioned by the Dutch government. According to the report, human embryo research is acceptable only under conditions, including the condition that “it must not be possible to obtain the envisaged results in any other way (through animal research or cell culture, for example)” (Health Council, 1998, p. 58). Further examples can be found in professional guidance documents. For example, the International Society for Stem Cell Research (ISSCR) states that human embryo research proposals “should include a discussion of alternative methods and provide a rationale for [...] performing the experiments in a human rather than animal model system” (ISSCR, 2016, p. 6).

In many European jurisdictions, such as the Netherlands, Belgium, the United Kingdom and France, the UAF-rule is enshrined in the law. For example, following the Dutch embryo law, it is permissible to use embryos for research only when there is “no alternative research method” (CBO, 2003), and in the UK, research on human embryos is only permitted when the Human Fertilisation and Embryology Authority (HFEA) “is satisfied that any proposed use of embryos [...] is necessary for the purposes of the research” (Human Fertilisation and Embryology Act, 1990, p. 60). In order to receive a human embryo research license from the HFEA, “the scientists applying for a research licence are asked to provide a lot of information about the research they wish to carry out and why the research could not be done using animal embryos or other types of cells” (HFEA, 2017).

Ideal chain of ART research steps

More specifically, in guidance pertaining to preclinical ART research, the UAF-rule translates into a preferred order of research steps in which animal studies precede human embryo research. For instance, according to the Health Council of the Netherlands, the ideal order of research steps involved in the responsible introduction of new ARTs is as follows: “animal-based experimental research, preclinical research using embryos, randomized trials with patient groups, and follow-up research among the children” (Health Council, 1998, p. 70). This same order of consecutive steps can be found in the recommendations of ESHRE’s Task Force Ethics and Law (Pennings et. al., 2007) and more recently in papers of a group of leading ART-scientists, presenting an “ideal paradigm”, or “pathway” from basic physiological to clinical research through intermediate steps of animal and human embryo research (Brison et. al., 2013; Harper et. al., 2012; Harper et. al., 2017; Mulder et. al., 2018). The adjective ‘ideal’ reflects that in practice this order cannot always be followed. For instance, when a useful animal model is lacking, or when a question cannot be answered in research using human embryos, these steps will have to be skipped. Inevitably that will make it a greater challenge to determine if the step towards first clinical applications can safely be made.

3. Implications for scientific practice

From a purely scientific point of view, there is no such thing as a preferred order of research-steps other than in terms of always choosing the optimal way of answering a given research question. In preclinical ART research, this may involve the use of different kinds of animal models, but also of human embryos, depending on what would best fit the precise aims of the study. There is not necessarily a conflict between this scientific perspective and the UAF-rule. For example, for many preclinical research questions (e.g. whether a new ART may lead to successful pregnancies in mammals, or to study transgenerational safety effects of new ARTs), human embryos are not suitable to begin with. It is clear that these questions can only be studied in animal models (Mulder *et. al.*, 2018; Sharpe 2018). Where human embryos are suitable (e.g. for studying the possible effect of a new ART on gene expression patterns during embryonic development), doing the same research with animal embryos would not be as informative, given the translational distance between those species and humans. As in such a case the same question cannot be answered in research using animal embryos, the UAF-rule does not stand in the way of directly moving to human embryo research.

The notion of an 'ideal chain' reflects however a more extensive interpretation than the above strict reading of the general rule. According to this interpretation, human embryo research should only be considered after promising results have first been obtained in animals, or when it is clear that animal studies cannot contribute at all to answering the question. When following this interpretation, research ethics committees will not allow scientists to directly move to human embryo research as long as prior animal studies might still lead to results that are at least partially relevant to answering the research question. Again, this is not necessarily at odds with the more pragmatic perspective of science, which always requires balancing the pros and cons of different approaches, including issues such as costs, uniformity and easy availability in large numbers. However, deciding this on a case by case basis is not possible when research ethics committees take the line that animal research should always precede the use of human embryos regardless of how useful that prior step would be. Our impression is that researchers in the field of ART often face difficulties receiving permission for human embryo research when data from prior animal studies, more specifically: studies using animal embryos, are lacking. This is ethically

problematic in so far as it may lead to a waste of valuable research time and money, but also because it violates a core principle of the ethics of animal research, namely that the inevitable impact on animal wellbeing should be proportional to the importance of the benefits to be gained through that research (Jans et. al., 2018).

4. The morality of interests

Apart from the quoted line in the Warnock report, we did not come across any specific ethical justifications for the UAF-rule. It seems to be taken for granted that using human embryos as research material is morally more problematic than using animals and should therefore only be considered when (further) studies using an animal model will not contribute to the answer of the research question. But how strong is the argument for this position? In this section, we consider what animals stand to lose when being used as research material and ask how the fate of human embryos compares to this.

The strong consensus acknowledges that, in contrast to mere tissues and cells, animals have a moral standing, which means that we are obliged to consider their needs and interests when making decisions affecting them (Warren, 1997). Philosopher McMahan describes this as the 'morality of interests' (MoI), which refers to how the interests of beings at the receiving end of our actions constrain what we may do to them (McMahan, 2002). This view is reflected in the European animal research legislation, which incorporates the Three Rs; Replace, Reduce and Refine the use of animals for research (European Union, 2010).

The more the use of animals in research leads to pain, discomfort, a negative psychological impact, or an infringement of species-specific behavior, the greater the human benefits must be in order for the research in question to be considered proportional. Preclinical animal research for testing new ARTs has a relatively low impact on the wellbeing of animals, as compared to animal research in other contexts, such as certain forms of cancer or burns research (Jans et. al., 2018). Nevertheless, all animals in research suffer some extent of discomfort and/or pain. As preimplantation human embryos by contrast are not capable of experiencing pain or psychological distress, embryo research does not have an impact on their wellbeing. Choosing human embryos instead of animals (when both can be used for

answering the research question) would reduce the total negative impact on well-being and would in so far be morally preferable.

A reduced level of wellbeing is not the only way in which interests can be harmed; also premature death can do so and – at least in humans – taking a life is generally considered to be the greatest possible infringement on the victim's interests. In ART research, all animals (except nonhuman primates) will ultimately be painlessly killed. Following McMahan, the wrongness of killing depends on (1) the amount of good a being loses by dying and (2) the strength of its time-relative interest in continuing to live. This strength is dependent on whether, and if so how strongly, a being is psychologically connected to its future self. (McMahan, 2002). In this context, it is less wrong to kill lower animals (e.g. mice) than higher animals (e.g. non-human primates), as the latter would lose a more psychologically complex life and are more strongly connected to their future life because of their higher psychological features.

What about killing human embryos in or after research? Does this deprive them of a future life? Clearly, if left-over IVF-embryos are used, this is not a meaningful question, as those embryos are destined to perish already prior to being considered as research material (Pennings & van Steirteghem, 2004). However, what if human embryos are especially created to serve as research material? Could one argue that they are *ab initio* deprived of a more valuable future life than even the highest animal? This 'future like ours argument' has been applied to the ethics of abortion by philosopher Don Marquis (Marquis, 1989). But whatever its worth in that context, the argument cannot simply be extended to make the case that preimplantation embryos would be deprived of 'a future like ours'. Such deprivation would require a level of connection of a present self to a future self that goes beyond what can be attributed to an embryo. We conclude that the MoI does not support the UAF-rule. In fact, it would seem to rather support the opposite.

5. The morality of respect

A second possible line of argument is that a special moral status can be attributed to human embryos regardless of whether they can be said to have interests. A reference to this special

moral status can, for example, be found in the quoted argument from the Warnock report (Warnock, 1985).

An argument to support this claim would have to explain what morally distinguishes human embryos from animals. This cannot be the simple fact that they are 'human', as that would amount to a speciesist fallacy. If the argument can be made, it must somehow connect with *why* we think human beings have a special moral status and then, secondly, explain what this entails for human embryos. The main non-religious answer to the first part of this question is that we consider (most) humans to be 'persons' (DeGrazia, 1997; Harris, 1999; Warren, 1997), which means that they have a high 'enough' degree of personhood properties, such as "moral agency, autonomy, the capacity for intentional action, rationality, self-awareness, sociability, and linguistic ability" (DeGrazia, 1997; 2008, p. 193).

McMahan speaks of what we owe to persons in terms of the 'Morality of Respect' (MoR). Even if some animals (great apes) may perhaps qualify as borderline persons (DeGrazia, 1997), animals generally fall below what he calls the 'person threshold', which means that they are not entitled to the same degree of protection as human individuals. Whereas the use of persons as 'mere means' is morally unacceptable, this does not hold for beings below this threshold: using them for entirely instrumental purposes can be acceptable under conditions (Jans et. al., 2018).

What does this mean for a possible special moral status of early human embryos? Only for those who regard these as persons does their moral status have the weight accorded by the MoR. Most people, however, accord a more limited moral status to the human embryo. This is often argued for in terms of those embryos having the capacity to develop into human beings who will then be persons. Clearly, this is not enough to grant them the full protection of the MoR, given that even as 'potential persons', they are not persons yet. In fact, the widely endorsed understanding is that in comparison with that of persons, the special moral status of human embryos and fetuses is initially quite low and only increases with *in utero* development. Still, this is in comparison with persons, not with animals, to whom such a special moral status is not ascribed at all (apart perhaps from great apes as borderline persons (DeGrazia, 1997)). The reference to "an added measure of respect" in the Warnock report suggests that it is on this difference between human embryos and animals that the

UAF-rule rests. However, it is silent about the opposite conclusion that seems to follow from the MoI.

6. The interests of oocytes donors

For preclinical safety and efficacy studies in the context of introducing new ARTs, it will often not be possible to do the research with spare embryos that are left over from IVF treatment. In those cases, research embryos will need to be especially created for the purpose, which requires the use of donated human oocytes. As oocyte donation for the benefit of science is a morally sensitive practice, this might be regarded as an argument for the UAF-rule, at least in cases where the use of spare embryos is not an option. In order to produce mature oocytes, the donor must undergo hormone stimulation and oocyte pick-up. The procedure is burdensome and may entail physical risks, including a risk of the woman developing ovarian hyper stimulation syndrome (OHSS). However, with careful protocols these risks can almost completely be avoided (Devroey et. al., 2011). Still, in the light of the demanding nature of the procedure, not many women are willing to volunteer. Indeed, some commentators have argued that the creation of embryos for the purpose of research can increase the risk of coercion for women (Baylis, 2000) or might threaten women's autonomy (Gerrand, 1993). Financial compensation schemes or so-called egg-sharing procedures (in which IVF-patients are offered free or reduced-cost treatment in return for donating part of the mature oocytes obtained in the process (Blyth, 2002)) have been criticized for inviting exploitation of vulnerable women. Here again, however, these risks can be minimized with carefully designed practices of information, counseling and consent (Mertes & Pennings, 2006). Although it is certainly true that the challenges of responsibly obtaining mature human oocytes for preclinical embryo research are an argument for using spare embryos whenever the research question can be answered with these as well, we think it is difficult to maintain that these challenges are so large as to settle the issue in favor of the UAF-rule. Moreover, it is expected that with new sources of research oocytes, including oocytes left over from cryostorage for fertility preservation (Mertes et.al., 2012), immature oocytes grown and matured in vitro (McLaughlin et. al., 2018), and stem cell derived oocytes (Segers et. al., 2017), any remaining moral challenges will eventually be overcome.

7. Conclusion and recommendations

Apparently, the UAF-rule is based on the view that human embryos have a special moral status, without including the MoI perspective. When both perspectives are included, it is less clear that the UAF-rule is morally tenable. In light of this, different scenarios can be considered. When doing so, we leave aside any possible legal barriers to using or creating human embryos for research that exist in certain countries. Assuming that these research options are not categorically forbidden and therefore unavailable, let us first consider the scenario of preclinical research for which left-over IVF-embryos could be used. As said, those embryos are destined to perish already prior to being considered as research material. Using them for research only determines how they will perish, not whether they will. To maintain that it would still be morally preferable not to make use of those embryos, but to create and use animals instead (even if only as sources of gametes and embryos), strikes us as perverse. However, preclinical ART research cannot always be performed with left-over IVF embryos. As said, it will often be necessary to create embryos for the purpose. Would that change matter? If human embryos have a special moral status, they should not be created for research that can also be carried out with morally less offensive research material. But again: is that the case where animals are concerned that would also have to be created for the purpose? The least one should say, is that this is not obvious, not even with the possible support from the argument pertaining to the morality of asking women to act as oocyte donors. We conclude that the moral arguments to support the UAF-rule are not evidently weightier than arguments supporting the opposite view, according to which animals deserve more protection than human embryos.

As no overriding moral arguments supporting the UAF-rule were found, we wonder why this rule is so widely supported. Our hypothesis is that it is based on 'human prejudice', which in this case invites the view that human embryos deserve more respect than even the highest animal, solely because they belong to the 'human species' (Savulescu & Bostrom, 2009). Obviously, this is no valid argument to uphold the UAF-rule.

We recommend that in legal documents and professional guidance, the UAF-rule is abandoned and replaced by a justificatory framework that does not a priori favor animal studies over the use of human embryos in research. The notion of a chain of ART research

steps should be modified accordingly. A core message must remain that the step to the clinic can only responsibly be made after adequate preclinical safety studies. Clearly, this should include animal studies aimed at charting potential health effects of new reproductive technologies at multiple stages of offspring development in more than one generation (Mulder et. al., 2018, Sharpe 2018). However, it should also include early developmental (e.g. gene expression) research done directly on human embryos that would bypass the uncertainties that go with animal-based studies (Brison, et. al., 2013). With potentially risky reproductive technologies on the horizon (such as reproduction using stem-cell derived gametes, or genetically modified gametes or embryos), it is important that optimal use can be made of either type of morally sensitive research material. As the ill-founded notion of the greater moral sensitivity of human embryo research may stand in the way of this, it should be abandoned so as to avoid waste and increase the relevance of preclinical safety research. In line with this, research ethics committees assessing protocols for the use of human embryos in preclinical ART research should refrain from requiring prior animal studies that are not expected to be informative.

References

- Baylis, F. (2000). Our cells/ourselves: creating human embryos for stem cell research. *Women's Health Issues*, 10(3), 140-145.
- Blyth, E. (2002). Subsidized IVF: the development of 'egg sharing' in the UK. *Human Reproduction*, 17(12), 3254-3259.
- Brison, D. R., Roberts, S. A., & Kimber, S. J. (2013). How should we assess the safety of IVF technologies?. *Reproductive biomedicine online*, 27(6), 710-721.
- CBO. (2003). Modelreglement Embryowet. In Gezondheidszorg Kvd (ed). Van Zuiden Communications. Alphen aan de Rijn.
- DeGrazia, D. (1997). Great apes, dolphins, and the concept of personhood. *The Southern journal of philosophy*, 35(3), 301-320.
- DeGrazia, D. (2008). Moral status as a matter of degree?. *The Southern Journal of Philosophy*, 46(2), 181-198.
- Devroey, P., Polyzos, N. P., & Blockeel, C. (2011). An OHSS-Free Clinic by segmentation of IVF treatment. *Human reproduction*, 26(10), 2593-2597.
- Dondorp, W., & de Wert, G. (2011). Innovative reproductive technologies: risks and responsibilities. *Human Reproduction*, 26(7), 1604-1608.
- European Union. (2010) *Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes*. Official Journal of the European Communities 276(33): 33-79
- Gerrand, N. (1993). Creating embryos for research. *Journal of applied philosophy*, 10(2), 175-187.
- Harper J, Jackson E, Sermon K, Aitken RJ, Harbottle S, Mocanu E, Hardarson T, Mathur R, Viville S, Vail A. (2017). Adjuncts in the IVF laboratory: where is the evidence for 'add-on' interventions? *Human Reproduction*, 32(3): 485-491.
- Harper, J., Cristina Magli, M., Lundin, K., Barratt, C. L., & Brison, D. (2012). When and how should new technology be introduced into the IVF laboratory?. *Human reproduction*, 27(2), 303-313.
- Harris, J. (1999). The concept of the person and the value of life. In *Personhood and Health Care* (pp. 99-111). Springer, Dordrecht..

- Health Council of the Netherlands. (1998) *IVF-related research*. Rijswijk: Health Council of the Netherlands: 1998/08E
- HFEA. <http://hfeaarchive.uksouth.cloudapp.azure.com/www.hfea.gov.uk/161.html> (20 November 2017, date last accessed).
- The Human Fertilisation and Embryology Act. 1990.
- ISSCR. (2016). *Guidelines for Stem Cell Research and Clinical Translation*.
- Jans, V., Dondorp, W., Goossens, E., Mertes, H., Pennings, G., & de Wert, G. (2018). Balancing animal welfare and assisted reproduction: ethics of preclinical animal research for testing new reproductive technologies. *Medicine, Health Care and Philosophy*, 21(4), 537-545.
- Marquis, D. (1989). Why abortion is immoral. *The Journal of Philosophy*, 86(4), 183-202.
- McLaughlin, M., Albertini, D. F., Wallace, W. H. B., Anderson, R. A., & Telfer, E. E. (2018). Metaphase II oocytes from human unilaminar follicles grown in a multi-step culture system. *MHR: Basic science of reproductive medicine*, 24(3), 135-142.
- McMahan J. (2002). *The ethics of killing: Problems at the margins of life*. Oxford: Oxford University Press.
- Mertes, H., & Pennings, G. (2007). Oocyte donation for stem cell research. *Human Reproduction*, 22(3), 629-634.
- Mertes, H., Pennings, G., Dondorp, W., & de Wert, G. (2012). Implications of oocyte cryostorage for the practice of oocyte donation. *Human reproduction*, 27(10), 2886-2893.
- Mulder, C. L., Serrano, J. B., Catsburg, L. A., Roseboom, T. J., Repping, S., & van Pelt, A. M. (2018). A practical blueprint to systematically study life-long health consequences of novel medically assisted reproductive treatments. *Human Reproduction*, 33(5), 784-792.
- NC3Rs. <https://www.nc3rs.org.uk/the-3rs> (4 May 2018, date last accessed).
- Pennings, G., de Wert, G., Shenfield, F., Cohen, J., Tarlatzis, B., & Devroey, P. (2007). ESHRE Task Force on Ethics and Law 13: the welfare of the child in medically assisted reproduction. *Human Reproduction*, 22(10), 2585-2588.
- Pennings, G., & Van Steirteghem, A. (2004). The subsidiarity principle in the context of embryonic stem cell research. *Human Reproduction*, 19(5), 1060-1064.

- Provoost V, Tilleman K, D'Angelo A, De Sutter P, de Wert G, Nelen W, Pennings G, Shenfield F, Dondorp W. (2014). Beyond the dichotomy: a tool for distinguishing between experimental, innovative and established treatment. *Human Reproduction*, 29(3): 413-417.
- Roseboom, T. J. (2018). Developmental plasticity and its relevance to assisted human reproduction. *Human Reproduction*, 33(4), 546-552.
- Savulescu J, Bostrom N. (2009) *Human enhancement*. Oxford: Oxford University Press.
- Segers, S., Mertes, H., de Wert, G., Dondorp, W., & Pennings, G. (2017). Balancing ethical pros and cons of stem cell derived gametes. *Annals of biomedical engineering*, 45(7), 1620-1632.
- Sharpe R. (2018). Of mice and men: long-term safety of assisted reproduction treatments. *Human Reproduction*, 33(5): 793-796.
- Van Steirteghem A. (2008). What next for assisted reproductive technology? A plea for an evidence-based approach. *Human reproduction*, 23(12): 2615-2616.
- Warnock M. (1985). *A question of life: the Warnock report on human fertilisation and embryology*. Oxford: Blackwell.
- Warren MA. (1997). *Moral status: Obligations to persons and other living things*. Gloucestershire: Clarendon Press.

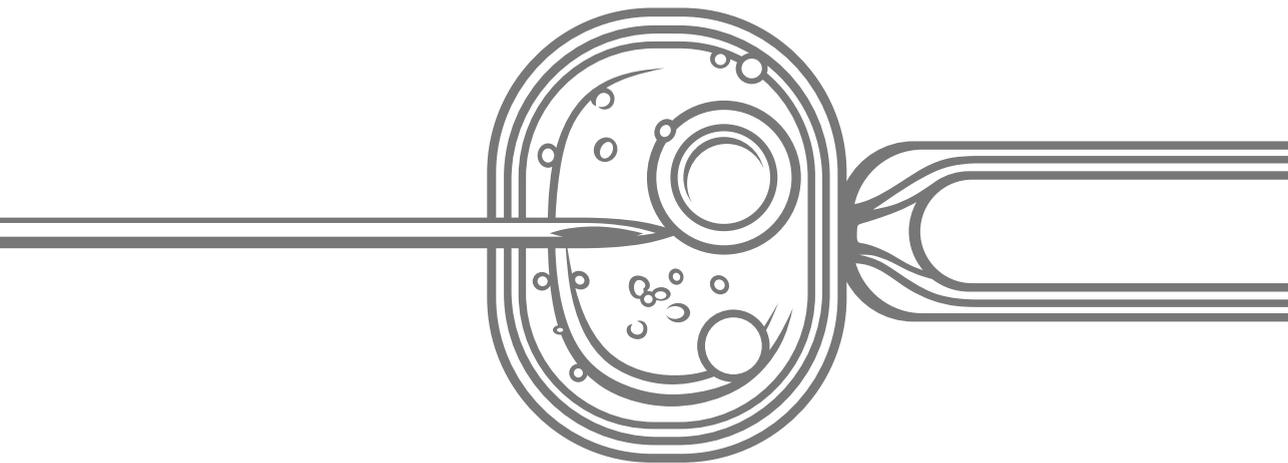


Chapter 4

Between innovation and precaution: How did offspring safety considerations play a role in strategies of introducing new reproductive techniques?

Published as:

Verna Jans, Wybo Dondorp, Sebastiaan Mastenbroek, Heidi Mertes, Guido Pennings, Hubert Smeets, Guido de Wert (2020). Between innovation and precaution: how did offspring safety considerations play a role in strategies of introducing new reproductive techniques? *Human Reproduction Open*.



1. Introduction

One concern within the wide range of ethical questions pertaining to developments in reproductive medicine, is whether treatment safety is sufficiently taken seriously when introducing new assisted reproductive technologies (ARTs) (Harper et. al., 2012, Mulder et. al., 2018, Pennings et. al., 2007, Provoost et. al., 2014, Sharpe, 2018, Van Steirteghem, 2008). Several commentators have pointed out that ARTs have often been, and are still being, introduced without systematic preclinical safety and effectiveness studies and without follow-up of children conceived with those technologies (Dondorp & de Wert, 2011, Harper et. al., 2017, Harper et. al., 2012, Sharpe, 2018). Although safety and effectiveness are different things, it is important to note that where evidence of effectiveness is lacking, there are no clear benefits that might render possible risks worth taking. In other words: to determine the proportionality of potentially risky new ARTs, data about safety and effectiveness are both needed.

This critique that these data have not systematically been sought, invites two different kind of questions. Empirically, the question is whether ‘responsible innovation’ has been a concern for the field and if so, how it has understood the practical implications of this idea for the development and introduction of potentially risky new ARTs. Ethically, the question is whether the field has indeed fallen short of its responsibilities in this respect and if so, how things can be improved.

In this paper, we aim to answer both questions. In the next three sections (2-4) we will first examine how offspring safety considerations did or did not play a role in the history of introducing new ARTs. We do so by presenting three case-studies: two from the past: intracytoplasmic sperm injection (ICSI) and preimplantation genetic testing (PGT), and one contemporary: Mitochondrial Replacement Therapy (MRT). Based on the findings, we will then move on to address the ethical question in the Discussion section (5).

2. Intracytoplasmic sperm injection

Technique

ICSI is a form of *in vitro* fertilization (IVF) whereby a single sperm cell is directly injected into an egg cell, which makes it possible for men with substantially low sperm quality to have

genetically related children. The first human births resulting from ICSI were reported in 1992 (Palermo et. al., 1992). Soon thereafter, the majority of IVF centers worldwide adopted the technique (Harper et. al., 2012). ICSI can be performed with ejaculated sperm (regular ICSI), but also with medically extracted sperm either from the epididymis (MESA) or from the testicles (TESE).

Offspring safety risks

The most relevant concerns about ICSI were twofold. Firstly, it was feared that the mechanical perforation of the egg cell might lead to oocyte injury or lysis (Laufer et. al., 1983). Secondly, several authors suggested that the sperm cells used in ICSI could have genetic abnormalities, because the technology involves bypassing the mechanism of natural selection. It was thought that this might affect the health of the offspring conceived through this technique (Cummins and Jequier, 1995, Cummins & Jequier, 1994). For MESA and TESE specifically, there were also health concerns due to the risk that the sperm cells used might be old or not fully matured (Cummins & Jequier, 1995, Küpker et. al., 2002, Martin-Deleon et. al., 1973). Older sperm cells show relatively high DNA damage, which might result in embryonic death or, in some cases, offspring abnormalities (Cummins and Jequier, 1995, Cummins and Jequier, 1994, Martin-Deleon et. al., 1973). The use of not fully matured sperm cells, potentially disrupting the genomic imprinting of the sperm cells, might lead to embryonic death, growth retardation, birth defects and functional disorders (Cummins & Jequier, 1995).

Safety considerations in strategies of introduction

Hardly any experimental studies were performed prior to the broad introduction of ICSI. Firstly, it was assumed at the time that testing ICSI in animal models was not technically feasible. Secondly, the accidentally discovered technology proved an immediate success in terms of helping couples with male infertility to have genetically related healthy children (Dondorp and de Wert, 2011). As a result, ICSI quickly proliferated in clinical practice with clinics often applying conditions such as avoiding the selection of immotile sperm in order to reduce possible risks. MESA and TESE soon became part of clinical practice in most countries as well. An exception to this picture was the decision of Dutch IVF professionals to withdraw

MESA and TESE from clinical application in view of the theoretical risks of these forms of ICSI, eventually leading to a national moratorium on MESA and TESE (Dutch Health Council, 1996), that was soon enacted into law. Lifting the moratorium would require animal studies providing a better understanding of the potential risks for the offspring. When reassuring data became available from animal research, sperm cell studies and clinical practice abroad, MESA (2000) and TESE (2007) were again allowed in the Netherlands, but only in a clinical research setting involving data-collection and follow-up (Kremer and Visser, 2008). The moratorium was fully lifted only in 2012 (for MESA) and 2014 (for TESE).

Although most commentators stressed the importance of follow-up research as part of the introduction of new ARTs, not many centers actually conducted follow-up of children conceived through ICSI. A notable exception to this was the pioneering center at Brussels Free University Hospital (Belva et. al., 2016, Belva et. al., 2011, Bonduelle et. al., 2004, Bonduelle et. al., 1994, Bonduelle et. al., 2005, Leunens et. al., 2007).

State of the art and current practice

It has since been established that ICSI does not lead to major abnormalities in offspring. (Leunens et. al., 2007, Pereira and Palermo, 2018). However, male subfertility has been associated with defects in the Y-chromosome. As these chromosomal defects are inherited by male offspring, it has been shown that boys conceived after ICSI more often also have lower sperm quality (Belva et. al., 2016). Additionally, ICSI is associated with a significantly increased risk of (very) preterm delivery (Wisborg et. al., 2010), lower mean birthweight (Bonduelle et. al., 2004) and an increased risk of rare imprinting disorders (Lazaraviciute et. al., 2014) in comparison to spontaneous conceived and non-IVF ART pregnancies. Yet, more studies are needed to investigate whether these specific health effects are caused by the technique itself or by other factors, such as underlying paternal conditions (Berntsen et. al., 2019, Devroey and Van Steirteghem, 2004). This is especially important in view of the fact that ICSI is increasingly offered for all causes of subfertility (Boulet et. al., 2015, Harris et. al., 2016, Nyboe Andersen et. al., 2008), albeit without data showing it to be more effective than conventional IVF for conditions other than male fertility problems (Li et. al., 2018, Van Rumste et. al., 2004). In the light of possible differences in the risk-profiles of both techniques,

this indication-creep is another instance of how safety and effectiveness considerations are both relevant for determining the proportionality of assisted reproduction.

3. Preimplantation Genetic Testing

Technique

Preimplantation genetic testing (PGT) enables the selection of embryos based on an evaluation of their genetic composition. PGT for monogenetic disorders (PGT-M), and PGT for structural rearrangements (PGT-SR) are indication-based applications aimed at helping at high risk couples to prevent transmitting a serious genetic disease to their offspring. These two types (historically referred to as preimplantation genetic diagnosis (PGD)) were introduced in the late 1980s (Handyside et. al., 1990). The third type, PGT for aneuploidy (PGT-A) was introduced to screen the chromosomal constitution of IVF-embryos with the aim of increasing the chances of achieving a successful pregnancy. This was historically called preimplantation genetic screening (PGS), and was introduced in 1995 (Verlinsky et. al., 1995).

The biopsy needed to perform PGT can take three different forms. Polar body biopsy is rarely used. It only offers information on the maternal contribution to the genome of the embryo. In blastomere biopsy (historically the most frequently used biopsy-type), one or two blastomeres are removed from a cleavage stage embryo on 3 days post-fertilization (~ 8 cells). In trophectoderm biopsy, several cells are removed from a blastocyst stage embryo on 5 days post-fertilization (150-300 cells).

Offspring safety risks

The main safety concern of PGT is the biopsy needed to collect cells for genetic testing. There are three possible risks associated with PGT biopsies for offspring thus conceived. First, there is the risk of the biopsy procedure itself. The use of a laser, additional actions outside an incubator or specific embryo culture media may have risks for the child conceived after PGT. For example, there is evidence that the use of specific embryo culture media affects important offspring parameters such as birth weight (Zandstra et. al., 2015).

Second, the removal of cells may have risks for the development of the embryo. Early studies in mice and humans have shown an association of blastomere biopsy with impaired implantation potential and reduced/delayed fetal development (Liu et. al., 1993, Tarín et. al., 1992, Tarín & Handyside, 1993). In recent years, the field is shifting to favoring trophectoderm biopsy over blastomere biopsy (Scott, et. al., 2013, Zacchini et. al., 2017). A main reason for this is the increased popularity of PGT-A in many countries, and the aim of overcoming its diagnostic difficulties due to mosaicism when using blastomere biopsy. Trophectoderm biopsy may seem less risky at first glance, as the biopsy does not affect the inner cell mass, of which the cells later form the fetus. In the past, it was thought that “the loss of a few mural trophectoderm cells should not be important embryologically, since this tissue does not seem to play a fundamental role in the later development of the embryo” (Edwards and Hollands, 1988). More recently, however, it is speculated that removing these cells could result in a smaller placenta, which could lead to incorrect embryo development and abnormalities in offspring (Zacchini et. al., 2017, Zhang et. al., 2019).

Third, for trophectoderm biopsy, the embryo needs longer in-vitro culturing than for blastomere biopsy (5 days instead of 3 days). The effect of extended embryo-culture is still under discussion and thus far little data is available (Youssef et. al., 2015). However, as the use of specific embryo culture media affects important offspring parameters such as birth weight (Zandstra et. al., 2015), longer embryo culturing may further increase these risks. Additionally, cryopreservation may be necessary more often. Singleton pregnancies following frozen embryo transfer are associated with higher risk of high birth weight babies compared with those from fresh embryo transfer (Maheshwari et. al., 2018a), although the interpretation of this finding has led to some debate (Somigliana et. al., 2018, Maheshwari et al, 2018b). Yet, data on cryopreservation in combination with PGT-A specifically is still limited (Penzias et. al., 2018).

Safety considerations in strategies of introduction

Prior to the introduction of PGT, preclinical research has been performed, particularly in the mouse model (Gardner, 1985, Monk et. al., 1987, Wilton et. al., 1989). After its introduction, research continued in mice, other animal models and in human embryos (Carson et. al., 1993,

Pierce et. al., 1997, Takeuchi et. al., 1992). Nevertheless, information on biopsy safety was limited (de Vos & van Steirteghem, 2001).

Only few safety measures aimed at decreasing possible offspring safety risks were taken when introducing the new technology. Some centers initially recommended their patients to consider prenatal diagnosis after PGT (Van der vorst et. al., 2000). This was however mainly to confirm the result of the diagnosis. One understanding at the time was that as biopsy damage would lead to failed implantation, there was no need for much concern regarding surviving embryos (Edwards & Hollands, 1988). Still, regulation involving limiting indications for what was then called PGD to serious conditions was partly also informed by the notion that possible biopsy risks should be proportional to the benefits of selective reproduction made possible by the procedure (de Wert et. al., 2014). Similar reasoning may have informed the 2001 decision by the UK Human Fertilization and Embryology Authority (HFEA) not to allow preimplantation tissue typing, stating that there were “risks arising from embryo biopsy and it was felt that the Ethics Committee had not taken proper account of the absence of evidence of no risk” (House of Commons, 2005). This ‘absence of evidence of no risk’ has not stood in the way of a rapid global proliferation of PGT-A. Even after its upgrade from PGS1 (blastomere biopsy and FISH) to PGS2 (trophectoderm biopsy and assessment of all chromosomes), evidence for the effectiveness of this procedure is still lacking and the treatment’s rationale is under discussion (Lawrenz et. al., 2019, Mastenbroek, 2013, Mastenbroek & Repping, 2014, Penzias et. al., 2018, Popovic et. al., 2018, Scott Jr et. al., 2012, Sermon et. al., 2016, van Loendersloot et. al., 2014). Furthermore, although concerns about the safety of blastomere biopsy (also referring to possible epigenetic effects) are recently stressed by authors with a stake in promoting PGT-A, there seems less interest in how the extended embryo culture needed for trophoblast biopsy may affect the relative risk-profiles of both biopsy procedures. Yet, it is well established that culture conditions have an impact on the epigenetics of developing embryos, which may well be more outspoken when the culture period is longer (Zandstra et. al., 2015). Only few clinical studies have compared blastomere biopsy with trophectoderm biopsy (Adler et. al., 2014, Coll et. al., 2018, McArthur et. al., 2008),

Only a few centers succeeded in conducting follow-up research on PGT, particularly in combination with blastomere biopsy. One of the leading groups was again the Free University

Brussels Hospital (Beukers et. al., 2013, Desmyttere et. al., 2009, Kuiper et. al., 2017, Liebaers et. al., 2009, Nekkebroeck et. al., 2008).

State of the art and current practice

Follow-up research does not show a higher degree of major abnormalities in children born after PGT in comparison to IVF or ICSI only (Banerjee et. al., 2008, Desmyttere et. al., 2009, Heijligers et. al., 2019, Heijligers et. al., 2018, Liebaers et. al., 2009, Nekkebroeck et. al., 2008). Although children born following PGT-M and PGT-SR in combination with blastomere biopsy do have an increased risk of adverse obstetric and neonatal outcomes, these seem mainly related to underlying parental conditions, except for an increased risk of placenta previa (Bay et. al., 2016). It is important to note that follow-up studies investigating the health of children born after PGT in combination with trophectoderm biopsy are still very limited.

PGT-A was originally recommended and carried out as an add-on to IVF for couples with advanced maternal age, repeated IVF failure, repeated miscarriage and severe male factor infertility. Regardless of the ongoing debate about its effectiveness, the technique is increasingly also offered to younger women with a good pregnancy prognosis (Mastenbroek & Repping, 2014, Penzias et. al., 2018).

4. Mitochondrial Replacement Therapy

Technique

Mitochondrial Replacement Therapy (MRT) is a form of *in vitro* fertilization, in which the future child's mitochondrial DNA comes from a donor. This technique has been developed to allow women who carry disease-causing variants in their mitochondrial DNA (mtDNA) to prevent passing these on to their offspring (Greenfield et. al., 2017, Hyslop et. al., 2016, Zhang et. al., 2017). There are several forms of MRT. In maternal spindle transfer (MST), the spindle-chromosome complex is transferred from the prospective mother's oocyte into a donor oocyte from which the same has been removed. The resulting oocyte will then be fertilized *in vitro*. In pronuclear transfer (PNT), the male and female pronuclei are removed from the recipient's fertilized egg prior to their fusing and are inserted into a fertilized donor egg from which the pronuclei are removed (Craven et. al., 2010).

Offspring safety risks

Apart from the risk of not preventing transmission of the disease, the main concern with MRT is that manipulation of the (fertilized) egg cell may have adverse effects for the resulting embryo (Adashi & Cohen, 2017). One study showed that “oocytes that had received their nucleus from a donor were less likely to develop into a blastocyst than oocytes who had had their own nucleus injected back into them” (Hyslop et. al., 2016). Yet, it is unclear whether this resulted from suboptimal interaction between the nucleus of the recipient and the mitochondria of the donor, or from the oocyte cryopreservation procedure (Eyre-Walker, 2017, Wei et. al., 2019).

Safety considerations in strategies of introduction

In comparison to ICSI and PGT, preclinical research on MRT was more extensive and more structured. The relevant studies were performed in animals (including nonhuman primates) and human embryos (Craven et. al., 2010, Paull et. al., 2013, Tachibana et. al., 2013, Tachibana et. al., 2009). The UK, in particular, has much invested in this research, as it was a precondition for the government to lift legal restrictions and for the HFEA to grant a clinical license. Although many research steps were aimed at improving efficiency, some of them also concerned improving safety, such as studies of epigenetic modifications and gene expressions in embryos derived from PNT and MST (HFEA, 2016).

Approaches of introducing MRT vary widely. In 2003, a collaboration between a Chinese group and a team from the US reported the first pronuclear transfer (PNT) (Zhang et. al., 2016). The attempt resulted in the death of two fetuses after selective fetal reduction. This ultimately led to the US and China imposing prohibitive regulatory policies, which still stand today (Ishii, 2018, Ishii & Hibino, 2018). In 2016, US-based scientists succeeded in delivering the first baby as a result of MST (Zhang et. al., 2017). The procedure was performed in a clinic in Mexico to avoid the prohibitory policy of the US. Since then, pregnancies and births using MRT have also been reported in Ukraine for cases of embryo arrest (PNT) and in Greece/Spain to overcome female infertility (MST), again making use of the lack of appropriate regulation (Ishii and Hibino, 2018). In 2015, the United Kingdom lifted legal restrictions banning the clinical use of MRT, however only for licensed clinics under strict conditions as formulated in scientific reviews conducted by the HFEA. For example, MRT

should only be offered to women who have a significant risk of transmitting a serious mitochondrial disease and where PGT is no alternative (HFEA, 2015). Additionally, in order to eliminate the risk of the disease reemerging in subsequent generations, it is recommended that any female born following MST or PST should be advised to use PGT if she wishes to have children of her own (HFEA, 2013). In 2017, the HFEA awarded the Newcastle group the first UK clinical license, which resulted in their first patient in 2018. This group has accepted various women with different indications, but so far no pregnancy has been reported. In line with HFEA recommendations, the Newcastle group aims to perform follow-up studies. However, whereas those recommendations insisted on long-term follow-up, for practical reasons the current aim is limited to gathering follow-up data until the age of 18 months only.

State of the art and current practice

As births following MRT are still very limited and no follow-up data is available, hardly any conclusions can be drawn regarding the effects in children thus conceived. Recently, several fertility centers worldwide have started offering MRT also as a means of 'rejuvenating' the eggs of older women with the promise of increasing their chances of a successful pregnancy, an idea probably inspired by ooplasmic transfer in the nineties (Cohen et. al., 1998). The recently reported life births after MRT in Greece and Ukraine were an instance of this (Devlin, 2019). This premature widening of indications has led the European Society of Human Reproduction & Embryology (ESHRE) to issue a strong condemnation, discouraging the use of MRT for fertility problems "until this technology has been proven to be effective and safe" (ESHRE, 2019). Although MRT for fertility problems may be less risky as carry-over is no issue here, the point remains that, due to a lack of evidence on effectiveness, there are no proven benefits to outbalance remaining risks.

5. Discussion

Our case studies provide important context for the increasingly louder claim that, until now, responsible innovation has not been a sufficiently high priority for the field. Where precisely and in what sense might the field have fallen short and, if so, what follows in terms of possible lessons for the future? To answer this, we make use of Per Sandin's four dimensions relevant

to dealing with uncertain risks. He presents these in the context of his discussion of the so called 'precautionary principle'. Following Sandin, the general structure of the precautionary principle can be reproduced as: "If there is (1) a threat, which is (2) uncertain, then (3) some kind of action (4) is mandatory" (Sandin, 1999, p. 891). Using these dimensions as a heuristic tool allows us to zoom in on the reasoning behind specific choices made with respect to introducing new ARTs.

It is important to note that using the four dimensions in this way, does not in itself entail a commitment to accepting the precautionary principle as a framework for introducing new ARTs. In fact, the so called 'Innovation Principle', where "the call is for the benefits of innovation to be weighed against known harm" (Read and O'Riordan, 2017, p. 12), can be understood as dealing with the same issues. The innovation principle provides, however, a different answer at the crucial juncture where the precautionary principle links the uncertainty dimension with a need to take action. By insisting that only *known* harm can be a ground for limiting the application of new technologies, the innovation principle emphasizes the importance of unhindered innovation, of which the benefits would be frustrated by unnecessary measures of precaution.

As our case studies suggest, it would be difficult to maintain that the field has failed to recognize that new ARTs, involving untested ways of handling human gametes and embryos, come with possible risks for children thus conceived. In each of the three cases studies, there has been quite some discussion of the fact that the then new technology in question (ICSI and its variants, PGT, MRT) might have adverse consequences for the health of the offspring of the applicants. In most cases, concerns about possible offspring risks related to the application of those technologies were hypothetical and had not (yet) been substantiated. Therefore, any account of how risky precisely a new technology was, had to include an emphasis on the uncertain nature of the threat, both regarding its chances of materializing and the seriousness of the possible impact this might have on the health of children thus conceived.

There is a range of possible ways of dealing with this uncertainty. On the one end of the spectrum, supporters of the innovation principle may reason that any measures to constrain the introduction of new ARTs or to limit their application for safety reasons should be based on evidence of a significant chance that serious harm could otherwise occur. This view might

invite moving new technologies to the clinic without costly and time-consuming preclinical safety studies and refraining from setting safety-related restrictions as long as children are born healthy and no calamities ensue. Some degree of such ‘innovation principle-reasoning’ may have contributed to the rapid introduction of ICSI in the absence of preclinical safety studies for which a suitable animal model was not available at the time.

On the other end of the spectrum, strong versions of the precautionary principle insist on the opposite: as long as new reproductive technologies pose an uncertain threat to the health of future children, they should not be introduced. Although we did not encounter explicit references to the precautionary principle in our case studies, the Dutch moratorium on clinical applications of MESA and TESE seems a good example of strong precautionary principle-reasoning, and perhaps the same can be said of the FDA’s decision to ban clinical research of MRT pending safety concerns.

If the innovation principle stands for giving innovation a green light in the face of uncertainty, and strong versions of the precautionary principle would stand for giving it a red light, intermediate, or weaker, forms of the precautionary principle may be understood as inviting an amber light policy, allowing innovation to ‘proceed with caution’. What characterizes these as forms of the precautionary principle, is their insistence that uncertainty should not be taken as a reason for refraining from safety measures (Sandin, 1999). However, in order to be proportional, these measures do not have to be prohibitive. Using examples from our case studies, relevant measures can, firstly, take the form of research aimed at reducing uncertainty both through preclinical studies using animals and/or human embryos (as in the MRT-case), and through systematic long-term follow-up (as done by the pioneering ICSI-center in Brussels). Secondly, further measures may include specific restrictions aimed at reducing avoidable risks in the application of the relevant technology. Examples from our case-studies include the requirement of using motile sperm for ICSI or setting strict indications for the first clinical applications. A more generally relevant consideration in this respect is that most couples seeking medically assisted reproduction are subfertile rather than infertile. This means that in many cases ‘tailored expectant management’ will help avoiding unnecessary MAR in those who do not need fertility treatment yet (Eijkemans et. al., 2017). Avoiding overtreatment is important not only for economic reasons, but also in view of reducing avoidable burdens and risks, including

offspring risks (Kersten et. al., 2015). Finally, a ‘proceed with caution’ approach would recommend against piling up procedures with uncertain risks, as in the suggestion at the time of introducing ICSI to use the then still new PGT technology for checking the genetic health of ICSI-embryos (Morris and Gleicher, 1996).

In each of our case studies, we found instances of expanding the use of a potentially risky technology to applications for which its effectiveness has not been proven. This applies to the increasing tendency of offering ICSI for non-male fertility problems, to the widespread use of PGT-A as a general add-on to IVF and to recent reports about MRT being offered for oocyte rejuvenation and embryonic arrest. Instead of taking safety as an absolute criterion, assessing whether a technique is safe enough is rather a matter of proportionality. By using a potentially risky technique for indications that lack evidence of effectiveness, those risks do not weigh up against the benefits. While it may be an open question whether the innovation principle could be invoked for justifying these practices (does the principle allow ignoring uncertain risks also if there are no proven benefits that might be frustrated by restrictive measures?), it is clear that these unproven expansions are squarely at odds with even the weakest versions of the precautionary principle.

Although our case-study analysis suggests that it would be wrong to conclude that responsible innovation has not been a concern of the field, it is also clear that this has not been a systematic and shared endeavor of the field as a whole. Relevant ESHRE guidance (Pennings et. al., 2007, Provoost et. al., 2014), seems to be taken by fertility centers as expressing non-committal suggestions rather than stating shared professional responsibilities. Clearly, this should not be a problem if the innovation principle is to determine what counts as responsible innovation. While that may sound quite attractive in a field that to a large extent is driven by commercial interests, it is not a tenable position from an ethical point of view. As acknowledged by ESHRE, professionals working in human reproduction inevitably assume a responsibility to take account of the welfare of the children they are causally involved in creating (Pennings et. al., 2007). As it seems that what we have referred to as weaker forms of precautionary principle are best able to capture this double responsibility, it is crucial that the corresponding approach to responsible innovation should not remain a pious wish. In terms of what Sandin refers to as the ‘command dimension’ of dealing with uncertain risks, our case-studies suggest that without the backing of legal

frameworks (such as those governing the introduction of MRT in the UK), current professional guidance lacks sufficient commanding power to ensure that new technologies are not introduced without research efforts aimed at reducing uncertainty about the nature and magnitude of offspring risks.

Inevitably, this leads to the question of what the threshold for acceptable risk should be. Clearly, the view that MAR is only acceptable if it leads to perfectly healthy children lays the bar too high. Not only does all reproduction come with risks, the trade-offs relevant to MAR are different from those in natural reproduction. However, the fact that some patients are desperate enough to accept even high risks for the child-to-be rather than remain childless does not mean that it should be left to individual patients or couples to determine what risks are still acceptable. As elsewhere in medicine – and even more so in MAR given professional co-responsibility for the welfare of the child – informed consent is a necessary, but not a sufficient condition for what might count as responsible treatment (Dondorp & de Wert, 2011). According to ESHRE's Task Force on Ethics and Law, the bottom line is that assisted reproduction should not entail a high risk that the child would have a seriously diminished quality of life (Pennings et. al., 2007). Still, more debate seems needed on the most appropriate criterion for acceptable risk in the context of introducing new ARTs.

6. Conclusion and recommendations

In each of our case studies, offspring safety concerns have been on the agenda of the debate about bringing the relevant technologies to the clinic. However, apart from settings where this was legally required, systematic safety and effectiveness studies (in line with ESHRE-guidance) were not always conducted. One possible reason for this is different views about how to deal with the uncertain (often theoretical) nature of the risk. As we have argued, an innovation principle-approach seems at odds with the notion that professionals in assisted reproduction have a responsibility to take account of the welfare of the children they are creating. While strong (prohibitive) forms of the precautionary principle seem difficult to justify in the light of the equally morally important interests of those dependent on medical help for having children, we suggest a policy of proceeding with systematic caution.

Although follow-up research has shown that up until now only few adverse health effects of new ARTs have emerged, this should not be taken as a reason for complacency. As new reproductive technologies are on the horizon that appear to come with potentially significant offspring risks (including the reproductive use of stem cell derived gametes or genetically modified embryos), this should be regarded as a matter of concern that may require both professional societies and governments to take a more active role in safeguarding the responsible development of these technologies. Although this should be primarily a matter of self-regulation, governments can play an important role in terms of providing a legal context for such self-regulation, and for ensuring compliance with its provisions.

Finally, more debate seems needed on determining the threshold for acceptable risk in the context of introducing new ARTs. Multiple stakeholders, including professional societies and patient organizations, should have a role in this.

References

- Adashi, E. Y., & Cohen, I. G. (2017). Ideology v. beneficence. *Journal of Law and the Biosciences*.
- Adler, A., Lee, H. L., McCulloh, D. H., Ampeloquio, E., Clarke-Williams, M., Wertz, B. H., & Grifo, J. (2014). Blastocyst culture selects for euploid embryos: comparison of blastomere and trophectoderm biopsies. *Reproductive biomedicine online*, 28(4), 485-491.
- Banerjee I, Shevlin M, Taranissi M, Thornhill A, Abdalla H, Ozturk O, Barnes J, Sutcliffe A. (2008). Health of children conceived after preimplantation genetic diagnosis: a preliminary outcome study. *Reproductive biomedicine online*, 16(3): 376-381.
- Bay B, Ingerslev HJ, Lemmen JG, Degn B, Rasmussen IA, Kesmodel US. (2016) Preimplantation genetic diagnosis: a national multicenter obstetric and neonatal follow-up study. *Fertility and sterility*, 106(6): 1363-1369.
- Belva, F., Bonduelle, M., Roelants, M., Michielsen, D., Van Steirteghem, A., Verheyen, G., & Tournaye, H. (2016). Semen quality of young adult ICSI offspring: the first results. *Human Reproduction*, 31(12), 2811-2820.
- Belva, F., Painter, R., Bonduelle, M., Roelants, M., Devroey, P., & De Schepper, J. (2012). Are ICSI adolescents at risk for increased adiposity?. *Human reproduction*, 27(1), 257-264.
- Berntsen, S., Söderström-Anttila, V., Wennerholm, U. B., Laivuori, H., Loft, A., Oldereid, N. B., Romundstad, L. B., Bergh, C., Pinborg, A. (2019). The health of children conceived by ART: 'the chicken or the egg?'. *Human reproduction update*, 25(2): 137-158.
- Beukers, F., van der Heide, M., Middelburg, K. J., Cobben, J. M., Mastenbroek, S., Breur, R., van der Lee, J. H., Hadders-Algra, M., Bos, A. F., & Kok, J. H. (2013). Morphologic abnormalities in 2-year-old children born after in vitro fertilization/intracytoplasmic sperm injection with preimplantation genetic screening: follow-up of a randomized controlled trial. *Fertility and sterility*, 99(2): 408-413.
- Bonduelle, M., Bergh, C., Niklasson, A., Palermo, G. D., & Wennerholm, U. B. (2004). Medical follow-up study of 5-year-old ICSI children. *Reproductive BioMedicine Online*, 9(1), 91-101.
- Bonduelle, M., Desmyttere, S., Buysse, A., Van Assche, E., Schietecatte, J., Devroey, P., Van Steirteghem, A., & Liebaers, I. (1994) Pregnancy: Prospective follow-up study of 55

- children born after subzonal insemination and intracytoplasmic sperm injection. *Human Reproduction*, 9(9): 1765-1769.
- Bonduelle, M., Wennerholm, U. B., Loft, A., Tarlatzis, B. C., Peters, C., Henriët, S., Mau, C., Victorin-Cederquist, A., Van Steirteghem, A., & Balaska, A. (2005). A multi-centre cohort study of the physical health of 5-year-old children conceived after intracytoplasmic sperm injection, in vitro fertilization and natural conception. *Human Reproduction*, 20(2): 413-419.
- Boulet, S. L., Mehta, A., Kissin, D. M., Warner, L., Kawwass, J. F., & Jamieson, D. J. (2015). Trends in use of and reproductive outcomes associated with intracytoplasmic sperm injection. *Jama*, 313(3), 255-263.
- Carson, S. A., Gentry, W. L., Smith, A. L., & Buster, J. E. (1993). Trophoctoderm micro biopsy in murine blastocysts: comparison of four methods. *Journal of assisted reproduction and genetics*, 10(6), 427-433.
- Cohen, J., Scott, R., Alikani, M., Schimmel, T., Munne, S., Levron, J., Wu, L., Brenner, C., Warner, C., & Willadsen, S. (1998). Ooplasmic transfer in mature human oocytes. *Molecular Human Reproduction*, 4(3): 269-280.
- Coll, L., Parriego, M., Boada, M., Devesa, M., Arroyo, G., Rodríguez, I., Coroleu, B., Vidal, F., & Veiga, A. (2018). Transition from blastomere to trophoctoderm biopsy: comparing two preimplantation genetic testing for aneuploidies strategies. *Zygote*, 26(3): 191-198.
- Craven, L., Tuppen, H. A., Greggains, G. D., Harbottle, S. J., Murphy, J. L., Cree, L. M., Murdoch, A. P., Chinnery, P. F., Taylor, R. W., & Lightowlers, R. N. (2010). Pronuclear transfer in human embryos to prevent transmission of mitochondrial DNA disease. *Nature*, 465(7294): 82.
- Cummins, J. & Jequier, A. M. (1995). Concerns and recommendations for intracytoplasmic sperm injection (ICSI) treatment. *Human Reproduction*, 10: 138-143.
- Cummins, J. M., & Jequier, A. M. (1994). Treating male infertility needs more clinical andrology, not less. *Human Reproduction*, 9(7), 1214-1219.
- De Vos, A., & Van Steirteghem, A. (2001). Aspects of biopsy procedures prior to preimplantation genetic diagnosis. *Prenatal diagnosis*, 21(9), 767-780.

- De Wert, G., Dondorp, W., Shenfield, F., Devroey, P., Tarlatzis, B., Barri, P., Diedrich, K., Provoost, V. & Pennings, G. (2014). ESHRE task force on ethics and Law22: preimplantation genetic diagnosis. *Human Reproduction*, 29(8): 1610-1617.
- Desmyttere, S., Bonduelle, M., Nekkebroeck, J., Roelants, M., Liebaers, I., & De Schepper, J. (2009). Growth and health outcome of 102 2-year-old children conceived after preimplantation genetic diagnosis or screening. *Early human development*, 85(12), 755-759.
- Devlin H. (2019). Baby with DNA from three people born in Greece. <https://www.theguardian.com/science/2019/apr/11/baby-with-dna-from-three-people-born-in-greece-ivf> (24 October 2019, date late accessed).
- Devroey, P., & Van Steirteghem, A. (2004). A review of ten years experience of ICSI. *Human Reproduction Update*, 10(1), 19-28.
- Dondorp, W., & de Wert, G. (2011). Innovative reproductive technologies: risks and responsibilities. *Human Reproduction*, 26(7), 1604-1608. Dutch Health Council. (1996). ICSI.
- Edwards, R. G., & Hollands, P. (1988). New advances in human embryology: implications of the preimplantation diagnosis of genetic disease. *Human Reproduction*, 3(4), 549-556.
- Eijkemans, M. J. C., Kersten, F. A. M., Lintsen, A. M. E., Hunault, C. C., Bouwmans, C. A. M., Roijen, L. H., Habbema, J. D. F. & Braat, D. D. M. (2017). Cost-effectiveness of 'immediate IVF' versus 'delayed IVF': a prospective study. *Human Reproduction*, 32(5): 999-1008
- ESHRE. (2019). Spindle transfer in the treatment of infertility: an ESHRE position statement. Accessed on 9 July 2019 at: <https://www.eshre.eu/Press-Room/ESHRE-News>.
- Eyre-Walker, A. (2017). Mitochondrial replacement therapy: are mito-nuclear interactions likely to be a problem?. *Genetics*, 205(4), 1365-1372.
- Gardner, R. L. (1985). Origin and development of the trophoctoderm and inner cell mass. In *Implantation of the Human Embryo: 2nd Bourne Hall Meeting* (pp. 155-178). Academic Press London.
- Gezondheidsraad: Commissie Herziening Planningsbesluit IVF. (1996). ICSI. Den Haag: gezondheidsraad, publikatie nr 1996/06. (Health Council of the Netherlands: Committee on In vitro fertilization. ICSI. The Hague. 1996. Publication number 1996/06).

- Greenfield, A., Braude, P., Flinter, F., Lovell-Badge, R., Ogilvie, C., & Perry, A. C. (2017). Assisted reproductive technologies to prevent human mitochondrial disease transmission. *Nature biotechnology*, 35(11), 1059.
- Handyside, A. H., Kontogianni, E. H., Hardy, K. R. M. L., & Winston, R. M. (1990). Pregnancies from biopsied human preimplantation embryos sexed by Y-specific DNA amplification. *Nature*, 344(6268), 768-770.
- Harper, J., Jackson, E., Sermon, K., Aitken, R. J., Harbottle, S., Mocanu, E., Hardarson, T., Mathur, R., Viville, S. & Vail, A. (2017). Adjuncts in the IVF laboratory: where is the evidence for 'add-on' interventions? *Human Reproduction*, 32(3): 488-491.
- Harper, J., Cristina Magli, M., Lundin, K., Barratt, C. L., & Brison, D. (2012). When and how should new technology be introduced into the IVF laboratory?. *Human reproduction*, 27(2), 303-313.
- Harris, K., Fitzgerald, O., Paul, R. C., Macaldowie, A., Lee, E., & Chambers, G. M. (2016). Assisted reproductive technology in Australia and New Zealand 2014. *Sydney: National Perinatal Epidemiology and Statistics Unit, the University of New South Wales*.
- Heijligers M., Peeters, A., van Montfoort, A., Nijsten, J., Janssen, E., Gunnewiek, F. K., de Rooy, R., van Golde, R., Coonen, E. & Meijer-Hoogeveen, M. (2019). Growth, health, and motor development of 5-year-old children born after preimplantation genetic diagnosis. *Fertility and sterility*, 111(6): 1151-1158.
- Heijligers, M., Verheijden, L., Jonkman, L., van der Sangen, M., Meijer-Hoogeveen, M., Arens, Y., van der Hoeven, M. & de Die-Smulders, C. (2018). The cognitive and socio-emotional development of 5-year-old children born after PGD. *Human Reproduction*, 33(11): 2150-2157.
- HFEA. (2013). Scientific review of the safety and efficacy of methods to avoid mitochondrial disease through assisted conception: update. Accessed on 5 November 2018 at: https://www.hfea.gov.uk/media/2612/mito-annex_viii-science_review_update.pdf.
- HFEA. (2015). The human fertilisation and embryology (mitochondrial donation) regulations. Accessed on 5 November 2018 at: <https://www.legislation.gov.uk/ukdsi/2015/9780111125816/contents>.
- HFEA. (2016). Scientific review of the safety and efficacy of methods to avoid mitochondrial disease through assisted conception: 2016 update. Accessed on 5 November 2018 at:

- https://www.hfea.gov.uk/media/2611/fourth_scientific_review_mitochondria_2016.pdf.
- House of Commons: Science and Technology Committee. (2005). Human Reproductive Technologies and the Law. Fifth Report of Session 2004–05. Volume II: Oral and written evidence. Accessed on 5 November 2018 at: <https://publications.parliament.uk/pa/cm200405/cmselect/cmsctech/7/7ii.pdf>.
- Hyslop, L. A., Blakeley, P., Craven, L., Richardson, J., Fogarty, N. M., Fragouli, E., Lamb, M., Wamaitha, S. E., Prathalingam, N. & Zhang, Q. (2016). Towards clinical application of pronuclear transfer to prevent mitochondrial DNA disease. *Nature*, 534(7607): 383-386.
- Ishii, T. (2018). Reproductive Medicine Involving Mitochondrial DNA Modification: Evolution, Legality, and Ethics. *European medical journal of reproductive health*, 4(1): 88-99.
- Ishii, T., & Hibino, Y. (2018). Mitochondrial manipulation in fertility clinics: regulation and responsibility. *Reproductive biomedicine & society online*, 5: 93-109.
- Kersten, F. A. M., Hermens, R. P. G. M., Braat, D. D. M., Hoek, A., Mol, B. W. J., Goddijn, M., & Nelen, W. L. D. M. (2015). Overtreatment in couples with unexplained infertility. *Human Reproduction*, 30(1), 71-80.
- Kremer, J. A., & Visser, H. (2008). Testicular sperm extraction (TESE) with intracytoplasmic sperm injection (ICSI) now allowed in the Netherlands. *Nederlands tijdschrift voor geneeskunde*, 152(3), 164-166.
- Kuiper, D., Bennema, A., la Bastide-van Gemert, S., Seggers, J., Schendelaar, P., Mastenbroek, S., Hoek, A., Heineman, M. J., Roseboom, T. J. & Kok, J. H. (2017). Developmental outcome of 9-year-old children born after PGS: follow-up of a randomized trial. *Human Reproduction*, 33(1): 147-155.
- Küpker, W., Al-Hasani, S., Johannisson, R., Sandmann, J., Ludwig, M., Jocham, D., & Diedrich, K. (2002). The use of cryopreserved mature and immature testicular spermatozoa for intracytoplasmic sperm injection: risks and limitations. *Seminars in reproductive medicine*, 20(1): 025-036
- Laufer, N., Pratt, B. M., DeCherney, A. H., Naftolin, F., Merino, M., & Markert, C. L. (1983). The in vivo and in vitro effects of clomiphene citrate on ovulation, fertilization, and

- development of cultured mouse oocytes. *American Journal of Obstetrics and Gynecology*, 147(6), 633-639.
- Lawrenz, B., El Khatib, I., Liñán, A., Bayram, A., Arnanz, A., Chopra, R., De Munck, N. & Fatemi, H. (2019). The clinicians' dilemma with mosaicism—an insight from inner cell mass biopsies. *Human Reproduction*, 34(6): 998-1010.
- Lazaraviciute, G., Kauser, M., Bhattacharya, S., Haggarty, P., & Bhattacharya, S. (2014). A systematic review and meta-analysis of DNA methylation levels and imprinting disorders in children conceived by IVF/ICSI compared with children conceived spontaneously. *Human reproduction update*, 20(6), 840-852.
- Leunens, L., Celestin-Westreich, S., Bonduelle, M., Liebaers, I., & Ponjaert-Kristoffersen, I. (2008). Follow-up of cognitive and motor development of 10-year-old singleton children born after ICSI compared with spontaneously conceived children. *Human Reproduction*, 23(1), 105-111.
- Li, Z., Wang, A., Bowman, M., Hammarberg, K., Farquhar, C., Johnson, L., Safi, N. & Sullivan, E. (2018). ICSI does not increase the cumulative live birth rate in non-male factor infertility. *Human Reproduction*, 33(7): 1322-1330.
- Liebaers, I., Desmyttere, S., Verpoest, W., De Rycke, M., Staessen, C., Sermon, K., Devroey, P., Haentjens, P. & Bonduelle, M. (2009). Report on a consecutive series of 581 children born after blastomere biopsy for preimplantation genetic diagnosis. *Human Reproduction*, 25(1): 275-282.
- Liu, J., Van den Abbeel, E., & Van Steirteghem, A. (1993). Preimplantation diagnosis: The in-vitro and in-vivo developmental potential of frozen and non-frozen biopsied 8-cell mouse embryos. *Human Reproduction*, 8(9), 1481-1486.
- Maheshwari, A., Pandey, S., Amalraj Raja, E., Shetty, A., Hamilton, M., & Bhattacharya, S. (2018a). Is frozen embryo transfer better for mothers and babies? Can cumulative meta-analysis provide a definitive answer?. *Human reproduction update*, 24(1), 35-58.
- Somigliana, E., Vanni, V. S., Busnelli, A., Reschini, M., Papaleo, E., & Viganò, P. (2018b). Excessive fetal growth in frozen embryo transfer: false alarm or clinical concern?. *Human reproduction update*, 24(4), 516-517.

- Martin-Deleon, P. A., Shaver, E. L., & Gammal, E. B. (1973). Chromosome abnormalities in rabbit blastocysts resulting from spermatozoa aged in the male tract. *Fertility and sterility*, 24(3), 212-219.
- Mastenbroek, S. (2013). One swallow does not make a summer. *Fertility and sterility*, 99(5), 1205-1206.
- Mastenbroek, S., & Repping, S. (2014). Preimplantation genetic screening: back to the future. *Human Reproduction*, 29(9), 1846-1850.
- McArthur, S. J., Leigh, D., Marshall, J. T., Gee, A. J., De Boer, K. A., & Jansen, R. P. S. (2008). Blastocyst trophectoderm biopsy and preimplantation genetic diagnosis for familial monogenic disorders and chromosomal translocations. *Prenatal Diagnosis: Published in Affiliation With the International Society for Prenatal Diagnosis*, 28(5), 434-442.
- Monk, M., Hardy, K., Handyside, A., & Whittingham, D. (1987). Preimplantation diagnosis of deficiency of hypoxanthine phosphoribosyl transferase in a mouse model for Lesch-Nyhan syndrome. *The Lancet*, 330(8556), 423-425.
- Morris, R. S., & Gleicher, N. (1996). Genetic abnormalities, male infertility, and ICSI. *Lancet*, 347(9011), 1277..
- Mulder, C. L., Serrano, J. B., Catsburg, L. A., Roseboom, T. J., Repping, S., & van Pelt, A. M. (2018). A practical blueprint to systematically study life-long health consequences of novel medically assisted reproductive treatments. *Human Reproduction*, 33(5), 784-792.
- Nekkebroeck, J., Bonduelle, M., Desmyttere, S., Van den Broeck, W., & Ponjaert-Kristoffersen, I. (2008). Mental and psychomotor development of 2-year-old children born after preimplantation genetic diagnosis/screening. *Human Reproduction*, 23(7), 1560-1566.
- Nyboe Andersen, A., Carlsen, E., & Loft, A. (2008). Trends in the use of intracytoplasmic sperm injection marked variability between countries. *Human reproduction update*, 14(6), 593-604.
- Palermo, G., Joris, H., Devroey, P., & Van Steirteghem, A. C. (1992). Pregnancies after intracytoplasmic injection of single spermatozoon into an oocyte. *The Lancet*, 340(8810), 17-18.

- Paull, D., Emmanuele, V., Weiss, K. A., Treff, N., Stewart, L., Hua, H., Zimmer, M., Kahler, D. J., Goland, R.S. & Noggle, S. A. (2013). Nuclear genome transfer in human oocytes eliminates mitochondrial DNA variants. *Nature*, 493(7437): 632.
- Pennings, G., de Wert, G., Shenfield, F., Cohen, J., Tarlatzis, B., & Devroey, P. (2007). ESHRE Task Force on Ethics and Law 13: the welfare of the child in medically assisted reproduction. *Human Reproduction*, 22(10), 2585-2588.
- Penzias, A., Bendikson, K., Butts, S., Coutifaris, C., Falcone, T., Fossum, G., Gitlin, S., Gracia, C., Hansen, K. & La Barbera, A. (2018). The use of preimplantation genetic testing for aneuploidy (PGT-A): a committee opinion. *Fertility and Sterility*, 109(3): 429-436.
- Pereira, N., & Palermo, G. D. (2018). Intracytoplasmic sperm injection: history, indications, technique, and safety. In *Intracytoplasmic Sperm Injection* (pp. 9-21). Springer, Cham.
- Pierce, K. E., Michalopoulos, J., Kiessling, A. A., Seibel, M. M., & Zilberstein, M. (1997). Preimplantation development of mouse and human embryos biopsied at cleavage stages using a modified displacement technique. *Human reproduction (Oxford, England)*, 12(2), 351-356.
- Popovic, M., Dheedene, A., Christodoulou, C., Taelman, J., Dhaenens, L., Van Nieuwerburgh, F., Deforce, D., Van den Abbeel, E., De Sutter, P. & Menten, B. (2018). Chromosomal mosaicism in human blastocysts: the ultimate challenge of preimplantation genetic testing? *Human Reproduction*, 33(7): 1342-1354.
- Provoost, V., Tilleman, K., D'Angelo, A., De Sutter, P., de Wert, G., Nelen, W., Pennings, G., Shenfield, F. & Dondorp, W. (2014). Beyond the dichotomy: a tool for distinguishing between experimental, innovative and established treatment. *Human Reproduction*, 29(3): 413-417.
- Read, R., & O'Riordan, T. (2017). The precautionary principle under fire. *Environment: Science and Policy for Sustainable Development*, 59(5), 4-15.
- Sandin, P. (1999). Dimensions of the precautionary principle. *Human and Ecological Risk Assessment: An International Journal*, 5(5), 889-907.
- Scott Jr, R. T., Ferry, K., Su, J., Tao, X., Scott, K., & Treff, N. R. (2012). Comprehensive chromosome screening is highly predictive of the reproductive potential of human embryos: a prospective, blinded, nonselection study. *Fertility and sterility*, 97(4), 870-875.

- Scott Jr, R. T., Upham, K. M., Forman, E. J., Hong, K. H., Scott, K. L., Taylor, D., Toa, X. & Treff, N. R. (2013). Blastocyst biopsy with comprehensive chromosome screening and fresh embryo transfer significantly increases in vitro fertilization implantation and delivery rates: a randomized controlled trial. *Fertility and sterility*, 100(3): 697-703.
- Sermon, K., Capalbo, A., Cohen, J., Coonen, E., De Rycke, M., De Vos, A., Delhanty, J., Fiorentino, F., Gleicher, N. & Griesinger, G. (2016). The why, the how and the when of PGS 2.0: current practices and expert opinions of fertility specialists, molecular biologists, and embryologists. *MHR: Basic science of reproductive medicine*, 22(8): 845-857.
- Sharpe, R. M. (2018). Of mice and men: long-term safety of assisted reproduction treatments. *Human Reproduction*, 33(5):793-796.
- Somigliana, E., Vanni, V. S., Busnelli, A., Reschini, M., Papaleo, E., & Viganò, P. (2018). Excessive fetal growth in frozen embryo transfer: false alarm or clinical concern?. *Human reproduction update*, 24(4), 516-517.
- Tachibana, M., Amato, P., Sparman, M., Woodward, J., Sanchis, D. M., Ma, H., Gutierrez, N. M., Tippner-Hedges, R., Kang, E. & Lee, H. S. (2013). Towards germline gene therapy of inherited mitochondrial diseases. *Nature*, 493(7437): 627-631.
- Tachibana, M., Sparman, M., Sritanaudomchai, H., Ma, H., Clepper, L., Woodward, J., Li, Y., Ramsey, C., Kolotushkina, O. & Mitalipov, S. (2009). Mitochondrial gene replacement in primate offspring and embryonic stem cells. *Nature*, 461(7262): 367-372.
- Takeuchi, K., Sandow, B. A., Morsy, M., Kaufmann, R. A., Beebe, S. J., & Hodgen, G. D. (1992). Preclinical models for human pre-embryo biopsy and genetic diagnosis. I. Efficiency and normalcy of mouse pre-embryo development after different biopsy techniques. *Fertility and sterility*, 57(2), 425-430.
- Tarín, J. J., Conaghan, J., Winston, R. M., & Handyside, A. H. (1992). Human embryo biopsy on the 2nd day after insemination for preimplantation diagnosis: removal of a quarter of embryo retards cleavage. *Fertility and sterility*, 58(5), 970-976.
- Tarín, J. J., & Handyside, A. H. (1993). Embryo biopsy strategies for preimplantation diagnosis. *Fertility and sterility*, 59(5), 943-952.
- Van Loendersloot, L., van Wely, M., van der Veen, F., Bossuyt, P., & Repping, S. (2014). Selection of embryos for transfer in IVF: ranking embryos based on their implantation

- potential using morphological scoring. *Reproductive biomedicine online*, 29(2), 222-230.
- Van Rumste, M. M. E., Evers, J. L. H., & Farquhar, C. M. (2004). ICSI versus conventional techniques for oocyte insemination during IVF in patients with non-male factor subfertility: a Cochrane review. *Human Reproduction*, 19(2), 223-227.
- Van Steirteghem, A. (2008). What next for assisted reproductive technology? A plea for an evidence-based approach. *Human reproduction*, 23(12): 2615-2616.
- Van der vorst, M., Staessen, C., Sermon, K., De Vos, A., Van de Velde, H., Van Assche, E., Bonduelle, M., Van der faellie, A., Lissens, W. & Tournaye, H. (2000). The Brussels' experience of more than 5 years of clinical preimplantation genetic diagnosis. *Human reproduction update*, 6(4): 364-373.
- Verlinsky, Y., Cieslak, J., Freidine, M., Lvakhnenko, V., Wolf, G., Kovalinskaya, L., White, M., Lifchez, A., Kaplan, B. & Moise, J. (1995). Pregnancies following pre-conception diagnosis of common aneuploidies by fluorescent in-situ hybridization. *MHR: Basic science of reproductive medicine*, 1(5): 265-269.
- Wei, W., Tuna, S., Keogh M, J., Smith, K. R., Aitman, T. J., Beales, P. L., Bennett, D. L., Gale, D. P., Bitner-Glindzicz, M. A & Black, G. C. (2019). Germline selection shapes human mitochondrial DNA diversity. *Science*, 364(6442):749.
- Wilton, L. J., Shaw, J. M., & Trounson, A. O. (1989). Successful single-cell biopsy and cryopreservation of preimplantation mouse embryos. *Fertility and sterility*, 51(3), 513-517.
- Wisborg, K., Ingerslev, H. J., & Henriksen, T. B. (2010). In vitro fertilization and preterm delivery, low birth weight, and admission to the neonatal intensive care unit: a prospective follow-up study. *Fertility and sterility*, 94(6), 2102-2106.
- Youssef, M. M., Mantikou, E., van Wely, M., Van der Veen, F., Al-Inany, H. G., Repping, S., & Mastenbroek, S. (2015). Culture media for human pre-implantation embryos in assisted reproductive technology cycles. *Cochrane Database of Systematic Reviews*.
- Zacchini, F., Arena, R., Abramik, A., & Ptak, G. E. (2017). Embryo biopsy and development: the known and the unknown. *Reproduction*, 154(5), 143-148.

- Zandstra, H., Van Montfoort, A. P., & Dumoulin, J. C. (2015). Does the type of culture medium used influence birthweight of children born after IVF?. *Human Reproduction*, *30*(3), 530-542.
- Zhang, J., Liu, H., Luo, S., Lu, Z., Chávez-Badiola, A., Liu, Z., Yang, M., Merhi, Z., Silber, S. J. & Munné, S. (2017). Live birth derived from oocyte spindle transfer to prevent mitochondrial disease. *Reproductive biomedicine online*, *34*(4): 361-368.
- Zhang, J., Zhuang, G., Zeng, Y., Grifo, J., Acosta, C., Shu, Y., & Liu, H. (2016). Pregnancy derived from human zygote pronuclear transfer in a patient who had arrested embryos after IVF. *Reproductive biomedicine online*, *33*(4), 529-533.
- Zhang, W. Y., von Versen-Höyneck, F., Kappahn, K. I., Fleischmann, R. R., Zhao, Q., & Baker, V. L. (2019). Maternal and neonatal outcomes associated with trophoctoderm biopsy. *Fertility and sterility*, *112*(2), 283-290.

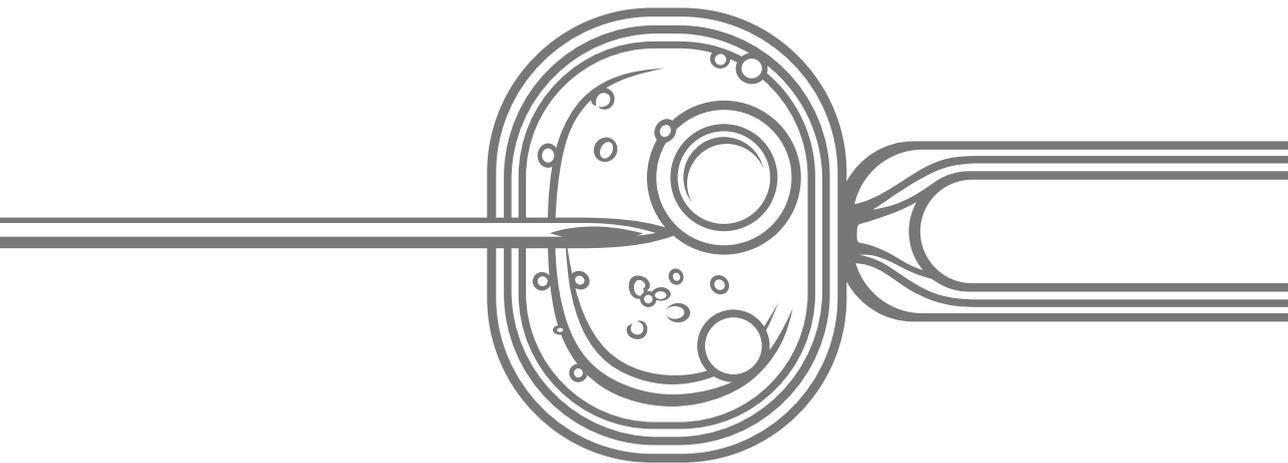


Chapter 5

Follow-up in the field of reproductive medicine: An ethical exploration

Published as:

Verna Jans, Wybo Dondorp, Maryse Bonduelle, Christine de Die, Heidi Mertes, Guido Pennings, Guido de Wert (2020). Follow-up in the field of reproductive medicine: An ethical exploration. *Reproductive BioMedicine Online*.



1. Introduction

Already more than ten years ago, the European Society of Human Reproduction & Embryology (ESHRE) proposed an ideal research trajectory for evaluating the safety of new assisted reproductive technologies (ARTs), consisting of preclinical research, clinical studies and follow-up (Pennings et. al., 2007). With potentially risky new ARTs on the horizon, such as germline gene editing or the use of stem cell derived gametes, and with long-term safety data still being scarce even for existing ARTs, this approach is increasingly seen as essential to responsible innovation in the field of medically assisted reproduction (MAR) (Harper, Magli, Lundin, Barratt, & Brison, 2012; Provoost et. al., 2014; Van Steirteghem, 2008).

In this paper, we will focus on the last step in ESHRE's recommended research trajectory: follow-up studies. Despite a general consensus on the importance of follow-up, still very few centers actually conduct it. Both practical and ethical concerns may hamper follow-up being performed responsibly. Without the illusion of being exhaustive, we will explore some of the most important ethical implications, issues and concerns related to performing follow-up studies of children born after ART, taking account of relevant experiences of the academic medical centers of Brussels (Free University), Belgium and Maastricht, the Netherlands.

In the next two sections we will chart the envisioned benefits of MAR follow-up for its various stakeholders, describe the types of procedures involved and hold current practice against ESHRE's normative framework recommending follow-up as a part of responsible innovation in MAR. In the second part of the paper, we discuss proportionality and respect for autonomy as the two main ethical issues relevant for responsibly conducting follow-up.

2. Background

Follow-up research in health care and medicine concerns the monitoring of patients' health over time after treatment. In the case of assisted reproduction, the subjects of such monitoring are not just patients (as for instance in the large number of studies that have been done to clarify whether hormonal stimulation puts women at a higher cancer risk later in their lives) but, importantly, also the resulting offspring. This article deals with follow-up of the latter.

Follow-up studies (both short and long-term) of children conceived after assisted reproduction generate health-impact data that may be important for four different groups of individual stakeholders, as well as for society at large. First, follow-up may benefit children born via the very procedures under evaluation. This is the case when health problems or health risks are found that are associated either with the use of a specific ART or its use in a specific population and that allow for timely treatment or prevention. For example, the inheritance of infertility is not yet fully understood. Should an increased risk for e.g. premature menopause be found in girls conceived through MAR in a specific population, preventive actions might be taken, such as oocyte vitrification to expand their reproductive options. Follow-up studies revealing conditions or health risks that are 'actionable' in the sense of allowing for timely treatment or prevention can be the basis for professional guidelines for screening or monitoring the children to whom this may apply.

Second, where long-term follow up becomes transgenerational by also looking at the health of the descendants of MAR-children, any actionable findings at that stage entail that the benefits of follow-up extend to those in this further generation, most of whom will themselves be conceived without medical help. The importance of transgenerational follow-up is widely acknowledged in relation to ARTs deliberately interfering with the human germline (National Academies of Sciences, 2016; Watts et. al., 2012), but seems also important for ARTs more generally, given that possible epigenetic effects may affect next generations.

Third, when findings are used to make MAR-procedures safer, this will benefit future children by reducing their chances of being born or having to live with ART-related malformations or disorders. Obviously, the parents and families of those children will benefit from such adapted procedures as well. For example, research has shown that multiple pregnancies are associated with a higher rate of prematurity and low birth weight, which puts the children at a significant risk of adverse health effects (Lu, Wang, & Jin, 2013). Therefore, ESHRE recommends single embryo transfer to avoid multiple pregnancies and advises not to transfer more than two embryos per cycle (ESHRE et. al., 2016). Another example is the observation of a higher risk of malformations in children resulting from IVF involving oocytes with cytoplasmic smooth endoplasmic reticulum aggregates (SERa). In 2011, this finding led to the so-called Istanbul Consensus recommending that those oocytes

should not be used in IVF (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011). In the light of conflicting evidence and the lack of a causal explanation, this recommendation was later changed into advising a case by case approach (ESHRE Special Interest Group of Embryology and Alpha Scientists in Reproductive Medicine, 2017), with the authors of a recent review recommending a policy of giving embryos derived from oocytes with SERa a lower priority ranking for transfer (Ferreux et al., 2019). Given that disregarding a small but significant percentage of oocytes that may still lead to healthy children has obvious implications for the success-rate of IVF, the latter example is a nice illustration both of the complex trade-off that may exist between IVF-safety and efficiency, and of the importance of conducting follow-up studies aimed at generating the high-quality evidence required for sustainable policy-decisions. Moreover, the example shows that epistemic uncertainty need not stand in the way of defining precautionary policy-responses, at least as long as more conclusive evidence is lacking (Jans et. al., HROpen 2020).

Finally, follow-up contributes to informed decision making by couples or individuals coming for MAR, as it enables a better understanding of the balance of benefits and risks of a certain reproductive technique. For example, follow-up on ICSI for male infertility has shown that, as compared to young men who were naturally conceived, those born after ICSI had lower scores on sperm concentration, sperm count and sperm motility parameters (Belva et al., 2016, p. 2817). Such information is important when counseling couples who consider fertility treatment. To make an informed decision, they should be able to weigh the pros and cons.

For society at large, it is important to consider that children conceived through MAR form an increasingly substantial part of the population, which means that even small MAR-related health risks can have a large impact on public health. For example, multiple studies have suggested that MAR-children have an elevated risk of developing cardiovascular diseases later in life (Ceelen, van Weissenbruch, Vermeiden, van Leeuwen, & Delemarre-van de Waal, 2008; Scherrer et. al., 2012; Scott et. al., 2010; Wikstrand, Niklasson, Strömmland, & Hellström, 2008). If adapted procedures, for instance the use of different culture media, can reduce those risks, this may have considerable population level benefits in addition to giving ART-conceived individuals healthier lives.

Follow-up of children conceived through MAR entails subjecting them to medical and/or psychological tests that may be more or less invasive and are often recurrent. Medical tests often consist of assessment of possible (major or minor) malformations, measuring (birth)weight, growth, hearing, visual acuity, blood pressure and neurological examination, including testing movement, muscle tension, reflexes, large and fine motor skills, and eye-hand coordination. These tests are performed by a pediatrician or another health care professional. Also reviewing the child's personal health record and, when relevant, reviewing the child's hospital records is usually part of follow-up. In long-term follow-up, aiming at examining long-term treatment effects, medical tests may also consist of analyzing blood samples or saliva specimen and of investigating reproductive health by analyzing semen samples or examining female reproductive organs. Psychological tests may include analyzing (1) cognitive, motor, and language developments, (2) behavior problems and socio-emotional development, and (3) parent-child relationship (Heijligers et. al., 2018; Zhan et. al., 2013). In practice, such tests include IQ tests, interviews and questionnaires.

In order to make information gained through follow-up available for studies evaluating long-term (transgenerational) consequences, centers are recommended to facilitate a uniform data collection: a registry (Provoost et. al., 2014). Such registries “could help document the risk of the different ART for prospective children. Over time, this could provide the required evidence base that would allow regulators, professional societies and individual clinicians to better address the needs and concerns of prospective parents and to improve the risk communication process” (Roy, Dupras, & Ravitsky, 2017, p. 440). Extensive registries are essential, because associations between aspects of a technique and certain health effects may only come to light much later. For example, documentation on specific embryo culture media is necessary to investigate possible associations between certain culture media and important offspring parameters, such as birth weight (Zandstra, Van Montfoort, & Dumoulin, 2015). Importantly, the value of such registries increases if a good registry of a control population exists.

3. Follow-up in practice

To what extent is ESHRE's normative framework, specifically recommending follow-up studies, put into practice? Although follow-up is considered an important aspect of the responsible innovation of new reproductive treatments by commentators in the ART field, in practice few centers have experience in this regard (Brison, Roberts, & Kimber, 2013) (Jans et al., 2020). Particularly long-term follow up studies are rare (Mulder et al., 2018). Although some centers have contributed to follow-up studies extensively, such as at Brussels Free University Hospital, these centers are still the exception.

In many cases where follow-up was actually conducted, studies had considerable methodological limitations (Mulder et al., 2018; Sutcliffe & Ludwig, 2007). One of the main limitations is the complexity of interpreting the scientific data, caused by the use of different IVF methods, the lack of controls and the differences in clinical definitions of outcome measures. Additionally, standard protocols on data collection are lacking, which hampers collecting uniform data on health effects in children born after MAR. (Mulder et al., 2018, p. 785).

Practical difficulties are often mentioned as a reason behind the lack of (methodologically strong) follow-up research worldwide. Follow-up studies often face scarcity of money, which hampers having an adequate sample size or complicates conducting follow-up at all. Particularly long-term studies are costly, time consuming, and often face large dropout or lost to-follow-up rates (Barnhart, 2013, p. 300). In many countries, 'basic' scientific funding does often not compensate for performing sufficient follow-up research and alternative funding is difficult to find. With some notable exceptions, such as Denmark (Norrman et al., 2020), governments often refrain from financially supporting fertility centers to perform minimal follow-up. Participation rates are affected by logistical aspects, such as the travel distance to the center where measurements are made, especially when families are expected to return every so many years, and the time investment this entails for them. A further issue is that not all MAR parents do inform their children about how they were conceived, which not only limits the willingness of MAR-families to participate in follow-up research, but also raises privacy issues complicating careful recruitment for such studies (Soini et al., 2006, p. 609).

Apart from possible privacy concerns related to parental disclosure decisions, other ethical issues (for instance: the need for MAR-families to lead normal lives, or the fact that children have not consented to be conceived through MAR) are sometimes raised as a ready and unchecked justification for not having to conduct adequate follow-up.

4. Proportionality

As for all scientific research with human subjects, the two main ethical issues relevant to assessing the acceptability of MAR-follow up are proportionality and respect for autonomy. We will discuss these aspects in the present and next sections respectively. The proportionality criterion generally requires that the possible benefits to be obtained from the relevant study clearly outweigh the possible risks and burdens for those asked to participate. Applying this to MAR follow-up, we will first consider what possible risks and burdens participation in such research may entail. As for a large part MAR follow-up is research with children, we will then also need to discuss that according to internationally accepted guidance, more stringent proportionality criteria apply for studies involving minors as research subjects.

4.1. Burdens and risks of participating in MAR-follow-up studies

As follow-up of children, adolescents or adults born through MAR entails subjecting them to medical and psychological tests, follow-up inevitably comes with possible risks and burdens. These depend primarily on the type of medical and psychological tests, their frequency and the overall duration of the follow-up study. To the extent that medical tests consist of non-invasive check-ups, such as measuring weight or blood pressure, risks are absent and burdens minimal. Yet, some tests may be more invasive. Think of blood drawing and tissue typing, especially when the latter requires a biopsy. Another example is the examination of the normal development of female reproductive organs, which requires a gynecological examination. Naturally, in all cases, the preferred method for such examinations should be discussed with and adapted to the participants.

Moreover, follow-up may cause psychological burdens for the children and their families. To have one's reproductive organs investigated and to be asked questions related

to sexual development and behavior may cause psychological discomfort. Moreover, follow-up consisting of frequent medical visits over a longer time-period may be psychologically exhausting. MAR-families may prefer to get on with their lives without repeated reminders of a history of infertility treatment. Additionally, children may face psychological burdens from feeling different from other children. Although there is evidence that psychological burdens for children born after MAR are limited, these studies did not specifically investigate psychological burdens caused by follow-up (Ponjaert-Kristoffersen et. al., 2004). In addition to the burdens and risks associated with testing procedures, a further possibly harmful effect of participating in MAR follow-up studies is the potential of findings revealing a serious health risk, leading to knowledge that may be experienced, both by the child and its parents, as casting a shadow over their lives. While this may be outweighed by the benefit of timely intervention in the case of actionable findings (e.g. when a girl is found to be at a high risk of developing Hereditary Breast or Ovarian Cancer (HBOC)), this is not the case for non-actionable findings (e.g. a high risk of a neurodegenerative disorder for which no treatment or prevention options exist, such as Huntington's disease). Finally, the extent to which participating in follow-up is considered burdensome, may differ between families and individuals, also taking account of highly specific practical aspects such as traveling distance and related time investment.

4.2 MAR follow-up as research with minors

MAR follow-up is mainly research with minors. Although MAR follow-up is ideally long-term and will continue into adulthood, this typically extends a research trajectory that starts with including very young children. Moreover, where long-term follow-up becomes transgenerational, young children in the next generation will again become research subjects. According to internationally accepted guidance (Doek et. al., 2009; Council of Europe, 1997; Council for International Organizations of Medical Sciences), minors may be included in research that has the potential to benefit themselves on the condition that the risks are minimized and outweighed by the prospect of individual benefit. If there is no such potential, research with minors is only allowed if (apart from general research ethics requirements) both the following conditions are fulfilled: 1) the research question can only

be answered by conducting research with minors (criterion of group-relatedness) and 2) the burdens and risks for those minors are minimal.

For the purpose of our analysis, this guidance provides important context. Whereas it seems that MAR follow-up meets the criterion of group-relatedness (only research with children can give insight into the possible impact of being MAR-conceived on children's development), procedures involving more than minimal burdens and risks may only be considered when the child itself can benefit. As indicated earlier, this may be the case when follow-up studies lead to findings revealing health problems or health risks that allow for timely treatment or prevention. An evidence-based likelihood of specific findings in this category would turn MAR follow-up into a form of 'therapeutic research' with minors. However, as long as such outcomes are a mere theoretical possibility, it would seem difficult to take them as a ground for justifying procedures involving more than minimal burdens and risks.

Although more than minimal burdens and risks are acceptable in principle where non-therapeutic follow-up studies are conducted with competent adults able to provide informed consent (see below), adult MAR-offspring is typically included in follow-up only at a stage where the evidence base regarding the safety of the relevant ART is already quite robust. This means that the potential benefits for children, couples and public health to be expected from such studies will be smaller as compared to those of follow-up in the initial stages of the development of a new ART. Moreover, as the ART through which they were conceived may over the years have undergone major adaptations, it will not always be obvious that the results of long-term follow up are still directly relevant for current practice. Therefore, although follow-up with adults allows giving more weight to third-party benefits when balancing these with the burdens and risks of follow-up, the case for doing so tends to weaken with the passing of time.

What we see here is a tension for which there is no easy solution: whereas the need for comprehensive safety data is greatest at the experimental stages of a potentially risky new technology, the scope for generating such data through follow-up studies in children is limited by the requirement of minimal burdens and risks.

5. Respect for autonomy

If a particular follow-up study with either children or adults meets the proportionality requirement, the further criterion of respect for autonomy requires that families and individuals should only be recruited on the basis of adequate informed consent. After summarizing what this entails, we will discuss whether respect for autonomy allows centers or professionals to encourage MAR-families to participate in follow-up.

5.1 Informed consent for MAR follow-up

As stated by Faden and Beauchamp, in order for consent to be autonomous, it must not only be 1) intentional and 2) based on a proper understanding of what is at stake, but also 3) voluntary, in the sense of being made without controlling external influences (Faden & Beauchamp, 1986). As emphasized in this last condition, whether or not to consent to participating in research should not just be an informed, but importantly also an unconstrained and voluntary decision. Only then can such decisions be autonomous in the sense of being fully owned by those who make them.

The age at which a youngster can provide independent consent for participating in research differs between 16 and 18 in different jurisdictions. Whereas some countries (such as the Netherlands and Norway) maintain a dual consent procedure as from the age of 12 (with the informed consent of both the child and its parents needed), many other countries require the child to be asked to 'assent' to the proposed research on the basis of information adapted to their level of understanding, again with different ages (often around 6-8) as from which this would be needed (EMA, 2019). This notion of 'assent' ('affirmative agreement') has been criticized for being insufficiently clear, also in terms of its demarcation with the concept of consent (Baines, 2011). From an ethical point of view, it is important not only that children who are mature enough to make the relevant decisions on the basis of an adequate understanding of the nature of the research and its implications, are allowed to do so, but also that those who are not yet competent in this sense, should not be burdened with choices that are beyond their capacity for autonomous agency (Hein & Jöesaar, 2015). For them, the decision about participating in research should be made by their parents. However, should a not-yet competent child consistently resist being subjected to certain research procedures,

such resistance should be respected unless participating is clearly in his or her own interest. As we have argued, this is hardly ever the case in the context of MAR follow-up.

5.2. Encouraging participation in MAR follow-up

According to a joint expert document from ESHRE and the European Society of Human Genetics (ESHG) “parents should be encouraged to take part in follow-up studies of health and development of their offspring” (Soini et. al., 2006). From the perspective of MAR professionals, this claim is understandable: they need the help of these families to be able to fulfill their own responsibility in the matter. But how does such encouragement relate to the requirement that participation should be based on voluntary consent, as stated above? Doesn’t encouragement entail a form of mild moral pressure and if so, would that be morally problematic? We suggest that this depends on whether MAR-families (parents and children) can be said to have a certain responsibility to participate in follow-up. To the extent that they do, it can be argued that encouraging them may under conditions be acceptable. In the remainder of this section we discuss two possible grounds for ascribing such a responsibility to MAR-families: parental responsibility and reciprocity.

5.2.1 Parental responsibility

At this point, a distinction must be made between the situation where a potentially risky new ART is still experimental or innovative, and the situation where it has become established treatment. In situations of the former kind, parents can be said to have a *prima facie* responsibility towards their children to enroll them in follow-up research that may lead to findings revealing possible health problems or health risks that allow for timely treatment or prevention (Ishii, 2019). The qualifier ‘*prima facie*’ refers to the condition that the research in question must be proportional, also taking account of the burdens for the parents and their family. If so, reminding them, if necessary, of this responsibility is not at odds with respecting their autonomy, as long as their freedom to decide otherwise is respected. In cases where an experimental new ART is introduced in the context of formal research aimed at generating initial data about safety and effectiveness, professionals may even go a step further and present pre-treatment agreement to participation in follow-up as a condition of access (Chen, Pascale, Jackson, Szvetecz, & Cohen, 2016; Watts et. al., 2012). Clearly this can only pertain to

follow-up of incompetent minors rather than also to long term follow-up for which the consent of the then adolescent or adult has to be sought. Moreover, a pre-treatment parental agreement to this effect cannot be enforced, given that research participants or their legal representatives can always decide to retreat from a research project in which they participate (National Academies of Sciences, 2016). Still, obtaining the parents' commitment to the implications of the experimental nature of the procedure is important in view of the need to as soon as possible reduce the uncertainties about offspring safety that may still exist at this early developmental stage of a potentially risky new ART (Ishii, 2019; Watts et. al., 2012).

As a further condition in this regard, couples having experimental treatment in a formal research setting should be asked to commit themselves to telling their child how it was conceived. A survey-based study exploring the welfare of children born after experimental ooplasmic transfer conducted in the late 1990s, showed that almost none of the respondent families had told their children about the mode of conception. The researchers surmise that precisely the use of an experimental procedure may have been an important factor in this decision (Chen et. al., 2016).

5.2.2 Reciprocity

Clearly, the appeal to 'parental responsibility' loses most of its force when the new ART has become established treatment, meaning that there is sufficient data to conclude that there are no serious concerns about its immediate safety (Provoost et. al., 2014). However, there may be a further moral reason for participation in MAR follow-up research, including for the active participation in long-term follow up of those who themselves have been conceived through MAR. The general case for this argument has, somewhat provocatively, been presented by bioethicist John Harris (Harris, 2005). He argues along two lines. Invoking the rule of rescue, he says that by volunteering as research subjects we can save others from serious harm and that if we can do so without excessive cost to ourselves, we have a duty to respond accordingly. Invoking the principle of fairness, his second argument is that as members of modern Western societies we all considerably benefit from advances in healthcare, and that, therefore, we have a *prima facie* moral duty to reciprocate by contributing to medical research as the social practice that produces those very benefits.

Failing to contribute while continuing to profit would amount to a form of morally reprehensible 'free riding' (Harris, 2005). According to Harris, the same reasoning extends to parents or other legal representatives deciding on behalf of children or incompetent adults. They may assume that "that the person they are making decisions for is, or would wish to be, a moral person who wants to or is in any event obliged to discharge his or her moral duties" (Harris, 2005, p. 246).

Harris' argument has met with criticism. Commentators have pointed out that his arguments fail to make the case for the supposed moral duty to participate in medical research (Brassington, 2007, 2011; Shapshay & Pimple, 2007). With regard to the rule of rescue: the connection between participating in research and possibly saving human lives as a result is far too indirect to generate a similar call of moral duty as in the proverbial pond case where a child is drowning before one's eyes. Moreover, as a candidate research participant, one is hardly ever the only person around who might be able to respond. And while it may be true that fairness requires some form of reciprocation, it does not specify content and kind of what this should mean in practice. The most that can be argued is that all of us have an 'imperfect moral duty' to further the general good and that participating in scientific research is just one of many different ways in which we may choose to do so (Shapshay & Pimple, 2007). Still, the idea of participating in medical research as a possible way of fulfilling this imperfect duty means that deciding to do so is not entirely supererogatory in the sense of going beyond what morality might require us to consider. If so, it may well make a difference that our discussion is not about medical research in general but about follow-up research, where only those can be asked to participate who have undergone the very procedure that is now being evaluated in the interest of future patients who may also need it. While with regard to a possible contribution to medical research in general, a person need not have any specific reason for choosing this over other contributions that she might make to the general good, things may well be different when considering a request to participate in follow-up research. Here the fact of being better placed than most others to make this specific contribution to the general good may amount to a moral reason for at least seriously considering such a request.

5. Conclusion and recommendations

Adequate follow-up research is an essential condition for responsibly introducing and responsibly innovating new reproductive techniques. In light of the clear benefits of follow-up for various stakeholders, the reproductive field should put more effort in actively implementing ESHRE's normative framework considering the execution of MAR follow-up.

As the relevant studies inevitably involve burdens and risks, it is essential that its proportionality is assessed on a case-by-case basis. Where children are concerned who cannot yet themselves consent to participating, more than minimal burdens and risks are acceptable only where it is reasonable to expect that they themselves can benefit from taking part in the relevant studies. Follow-up research requires the active involvement of MAR families, with parents consenting for their children and adolescents asked to do so for themselves. When seeking their voluntary consent for participating in follow-up that meets the condition of proportionality, professionals may encourage MAR families to partake in follow-up research. The case for this is strongest where a potentially risky experimental ART is concerned and such encouragement refers to a parental responsibility for the welfare of the child to be. Where such a new ART is introduced in the context of formal research, pre-treatment agreement to participation in follow-up may be presented as a (non-enforceable) condition for access. This also requires a parental commitment to tell their children how they were conceived. Where the relevant ART has become established treatment, there may still be a ground for encouraging both MAR-parents and -children to partake in follow-up research based on the reasoning that they are better placed than others to make this specific contribution to the general good.

References

- Baines, P. (2011). Assent for children's participation in research is incoherent and wrong. *Archives of disease in childhood*, 96(10), 960-962.
- Barnhart, K. T. (2013). Assisted reproductive technologies and perinatal morbidity: interrogating the association. *Fertility and Sterility*, 99(2), 299-302.
- Belva, F., Bonduelle, M., Roelants, M., Michielsens, D., Van Steirteghem, A., Verheyen, G., & Tournaye, H. (2016). Semen quality of young adult ICSI offspring: the first results. *Human Reproduction*, 31(12), 2811-2820.
- Brassington, I. (2007). John Harris' argument for a duty to research. *Bioethics*, 21(3), 160-168.
- Brassington, I. (2011). Defending the duty to research? *Bioethics*, 25(1), 21-26.
- Brison, D. R., Roberts, S. A., & Kimber, S. J. (2013). How should we assess the safety of IVF technologies? *Reproductive Biomedicine Online*, 27(6), 710-721. doi:10.1016/j.rbmo.2013.09.006
- Ceelen, M., van Weissenbruch, M. M., Vermeiden, J. P., van Leeuwen, F. E., & Delemarre-van de Waal, H. A. (2008). Cardiometabolic differences in children born after in vitro fertilization: follow-up study. *The Journal of Clinical Endocrinology & Metabolism*, 93(5), 1682-1688.
- Chen, S. H., Pascale, C., Jackson, M., Szvetecz, M. A., & Cohen, J. (2016). A limited survey-based uncontrolled follow-up study of children born after ooplasmic transplantation in a single centre. *Reproductive Biomedicine Online*, 33(6), 737-744.
- Council of Europe. (1997). *Convention for the protection of human rights and dignity of the human being with regard to the application of biology and medicine: convention on human rights and biomedicine; Oviedo, 4. 4. 1997: Conseil de l'Europe, Service de l'Édition et de la Documentation.*
- Council for International Organizations of Medical Sciences. (2017). International ethical guidelines for health-related research involving humans. *International ethical guidelines for health-related research involving humans.*
- Doek, J., Breuker, K., Knibbe, C., Offringa, M., Passchier, J., & Steinkamp, N. (2009). Advies medisch-wetenschappelijk onderzoek met kinderen. *Den Haag.*

- EMA. (2019, 23 October 2019). *Working Group on Ethics of the European Network of Pediatric Clinical Trials in Europe 2015*
- ESHRE, De los Santos, M. J., Apter, S., Coticchio, G., Debrock, S., Lundin, K., . . . Verheyen, G. (2016). Revised guidelines for good practice in IVF laboratories (2015). *Human Reproduction*, *31*(4), 685-686.
- Faden, R. R., & Beauchamp, T. L. (1986). *A history and theory of informed consent*: Oxford University Press.
- Ferreux, L., Sallem, A., Chargui, A., Gille, A.-S., Bourdon, M., Maignien, C., . . . Pocate-Cheriet, K. (2019). Is it time to reconsider how to manage oocytes affected by smooth endoplasmic reticulum aggregates? *Human Reproduction*, *34*(4), 591-600.
- Harper, J., Magli, M. C., Lundin, K., Barratt, C. L., & Brison, D. (2012). When and how should new technology be introduced into the IVF laboratory? *Human Reproduction*, *27*(2), 303-313. doi:10.1093/humrep/der414
- Harris, J. (2005). Scientific research is a moral duty. *Journal of Medical Ethics*, *31*(4), 242-248.
- Heijligers, M., Verheijden, L., Jonkman, L., van der Sangen, M., Meijer-Hoogeveen, M., Arens, Y., . . . de Die-Smulders, C. (2018). The cognitive and socio-emotional development of 5-year-old children born after PGD. *Human Reproduction*, *33*(11), 2150-2157.
- Hein, V., & Jöesaar, H. (2015). How perceived autonomy support from adults and peer motivational climate are related with self-determined motivation among young athletes. *International Journal of Sport and Exercise Psychology*, *13*(3), 193-204.
- Ishii, T. (2019). Should Long-Term Follow-up Post-Mitochondrial Replacement be Left up to Physicians, Parents, or Offspring? *The New Bioethics*, *25*(4), 318-331.
- Verna Jans, Wybo Dondorp, Sebastiaan Mastenbroek, Heidi Mertes, Guido Pennings, Hubert Smeets, Guido de Wert (2020). Between innovation and precaution: how did offspring safety considerations play a role in strategies of introducing new reproductive techniques? *Human Reproduction Open*
- Lu, Y.-h., Wang, N., & Jin, F. (2013). Long-term follow-up of children conceived through assisted reproductive technology. *Journal of Zhejiang University Science B*, *14*(5), 359-371.

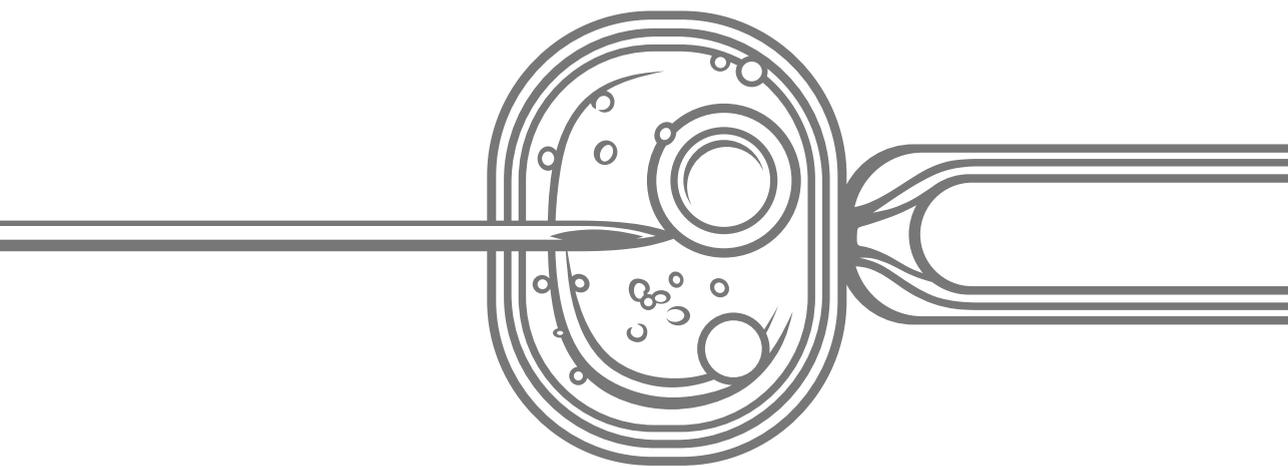
- Mulder, C. L., Serrano, J. B., Catsburg, L. A., Roseboom, T. J., Repping, S., & van Pelt, A. M. (2018). A practical blueprint to systematically study life-long health consequences of novel medically assisted reproductive treatments. *Human Reproduction*, *33*(5), 784-792.
- National Academies of Sciences, E., and Medicine. (2016). *Mitochondrial replacement techniques: Ethical, social, and policy considerations*. Washington, DC: The National Academies Press.
- Norrman, E., Petzold, M., Clausen, T., Henningsen, A., Opdahl, S., Pinborg, A., . . . Wennerholm, U. (2020). Type 1 diabetes in children born after assisted reproductive technology: a register-based national cohort study. *Human Reproduction*, *35*(1), 221-231.
- Pennings, G., de Wert, G., Shenfield, F., Cohen, J., Tarlatzis, B., & Devroey, P. (2007). ESHRE Task Force on Ethics and Law 13: the welfare of the child in medically assisted reproduction. *Human Reproduction*, *22*(10), 2585-2588. doi:10.1093/humrep/dem237
- Ponjaert-Kristoffersen, I., Tjus, T., Nekkebroeck, J., Squires, J., Verté, D., Heimann, M., . . . Wennerholm, U.-B. (2004). Psychological follow-up study of 5-year-old ICSI children. *Human Reproduction*, *19*(12), 2791-2797.
- Provoost, V., Tilleman, K., D'Angelo, A., De Sutter, P., de Wert, G., Nelen, W., . . . Dondorp, W. (2014). Beyond the dichotomy: a tool for distinguishing between experimental, innovative and established treatment. *Human Reproduction*, *29*(3), 413-417. doi:10.1093/humrep/det463
- Roy, M.-C., Dupras, C., & Ravitsky, V. (2017). The epigenetic effects of assisted reproductive technologies: ethical considerations. *Journal of developmental origins of health and disease*, *8*(4), 436-442.
- Scherrer, U., Rimoldi, S. F., Rexhaj, E., Stuber, T., Duplain, H., Garcin, S., . . . Allemann, Y. (2012). Systemic and pulmonary vascular dysfunction in children conceived by assisted reproductive technologies. *Circulation*, *125*(15), 1890-1896.
- Scott, K. A., Yamazaki, Y., Yamamoto, M., Lin, Y., Melhorn, S. J., Krause, E. G., . . . Tamashiro, K. L. (2010). Glucose parameters are altered in mouse offspring produced by assisted reproductive technologies and somatic cell nuclear transfer. *Biology of reproduction*, *83*(2), 220-227.

- Shapshay, S., & Pimple, K. D. (2007). Participation in biomedical research is an imperfect moral duty: a response to John Harris. *Journal of Medical Ethics*, 33(7), 414-417.
- Soini, S., Ibarreta, D., Anastasiadou, V., Aymé, S., Braga, S., Cornel, M., . . . Gianaroli, L. (2006). The interface between assisted reproductive technologies and genetics: technical, social, ethical and legal issues. *European Journal of Human Genetics*, 14(5), 588.
- Sutcliffe, A. G., & Ludwig, M. (2007). Outcome of assisted reproduction. *The Lancet*, 370(9584), 351-359.
- Van Steirteghem, A. (2008). What next for assisted reproductive technology? A plea for an evidence-based approach. *Human Reproduction*, 23(12), 2615-2616.
- Watts, G., Braude, P., Flinter, F., Harding, S., Lewens, T., & Parker, M. (2012). *Novel techniques for the prevention of mitochondrial DNA disorders: an ethical review*. Retrieved from London:
- Wikstrand, M. H., Niklasson, A., Ström, K., & Hellström, A. (2008). Abnormal vessel morphology in boys born after intracytoplasmic sperm injection. *Acta Paediatrica*, 97(11), 1512-1517.
- Zandstra, H., Van Montfoort, A. P., & Dumoulin, J. C. (2015). Does the type of culture medium used influence birthweight of children born after IVF? *Human Reproduction*, 30(3), 530-542.
- Zhan, Q.-t., Pan, P.-p., Xu, X.-r., Lou, H.-y., Lou, Y.-y., & Jin, F. (2013). An overview of studies on psychological well-being in children born following assisted reproductive technologies. *Journal of Zhejiang University Science B*, 14(11), 947-960.



Chapter 6

General discussion



Introduction

In the previous chapters, I explored how the field of MAR should responsibly deal with safety risks associated with introducing new ARTs. Using, and at the same time reflecting on, the normative framework of responsible innovation in MAR as proposed by ESHRE, I answered this question taking account of the consecutive steps of the proposed research trajectory: preclinical research (paper 1 and 2), clinical research (part of paper 3 and 4) and follow-up (paper 4). In this final chapter, I first summarize the main findings of my research. In the second part, I discuss the wider implications of my study and provide recommendations to adjust and refine ESHRE's normative framework.

1. Summary

Chapter two

As widely recognized, the field of MAR has often introduced new ARTs without sufficient research on safety and efficacy, thereby posing unnecessary risks on patients and ART-conceived offspring. Preclinical research, including animal research, has been recommended to introduce new ARTs more responsibly. Animal research can give insight into epigenetic, developmental and transgenerational risk of a technology in a relatively short time-period (Brison et. al., 2013), which may help to reduce these risks before applying the technology to humans.

Simultaneously with the acknowledgement that more preclinical research, including animal research, is needed in the ART field, there is growing international support for the Three Rs (Replace, Reduce and Refine the use of animals in research). The call for more animal research on the one hand and for avoiding animal research on the other, results in an ethical tension. This underlines the urgency of an explicit justification of the use of animals for the development and preclinical testing of new ARTs. Therefore, in the second chapter, I questioned whether the use of animals for efficacy and safety research on new ARTs can be ethically justified.

In order to answer this question, I used an ethical framework concerning animal research of considerable consensus, based on the premise that the moral status of animals is not comparable to that of humans (see also chapter 3). In contrast to humans, the use of

animals as a 'mere means' is considered acceptable under the principles of subsidiarity and proportionality (DeGrazia, 2003). The subsidiarity principle requires that the benefits to be derived from animal research cannot be obtained through other, less morally problematic, means. Subsequently, the proportionality principle requires that the level of infringement on animal wellbeing is proportional to the moral weight of those benefits. Both these principles have to be met in order for animal research to be morally acceptable.

It is implausible that alternative models, such as microorganisms (e.g. yeasts) and the use of animal/human *in vitro* cell and tissue cultures, can replace all ART animal research, because such models cannot provide information on the development of embryos and fetuses, nor on the health of (multigenerational) offspring. This means that ART animal research meets the principle of subsidiarity. Yet, a possibly important method to minimize the use of animals could be preclinical research on human embryos *in vitro*. This poses, however, an ethical dilemma considering the moral value of human embryos, which is addressed in the third chapter.

To assess whether the principle of proportionality can be met, the moral cost (impact of animal ART research on animal wellbeing) is balanced against the possible (moral) benefits of ART research. Potentially harmful procedures in animal ART research mainly consist of hormone stimulation injections or removal of tissue under sedation. At most, this may lead to short and light pain. Ultimately, most research animals will be euthanized when they are no longer useful for the research (except for e.g. nonhuman primates). Comparing this to other generally accepted animal research practices (i.e. cancer or burn research) and with the use of animals in the meat industry, the moral cost of animal ART research is relatively low. The moral benefit of ART research is assessed on two levels. On the first level, it is questioned whether the benefits of MAR *itself* are important enough to justify the use of animals. The main aim of MAR is to help people to have a genetically related child. As views on the importance of the genetic link vary widely, it is debatable whether MAR serves a vital human need in the same way as some other medical interventions do. It would, however, neither be correct to say that it merely serves a trivial preference. As those different views connect to the diverse understandings of what a flourishing life means, and in light of respecting that diversity in our liberal society, the practice of MAR does indeed serve a morally weighty aim. Comparing this aim to the widely accepted view that the use of animals

for food can be acceptable under conditions (that are obviously often violated in the mass-scale practice of the food industry) it seems difficult to maintain that research aimed at improving MAR through preclinical testing of ARTs would not be sufficiently important to allow the conditional use of animals for that purpose as well. Additionally, the earlier observation that, also in comparison with the use of animals in the meat industry, preclinical ART research has a relatively low impact on animal wellbeing, further supports this conclusion.

Moving on to the second level of the moral benefit of ART research, it is questioned whether the aims of ART *research* are important enough to justify the use of animals. Professionals providing assisted reproductive treatments have some level of co-responsibility for the welfare of their patients and the future children that they help bring into being (Dondorp & de Wert, 2011; G. Pennings et. al., 2007) Testing the efficacy of new ARTs prior to clinical introduction is important to avoid subjecting women (and sometimes men) to potentially risky treatment without sufficient benefit. Testing the safety is important to reduce the risk that children conceived after ART may be born with malformations or experience health problems during the rest of their lives. Given these morally important aims, balanced against the relatively low impact on animal wellbeing, the use of animals for testing the efficacy and safety of new ARTs is proportional.

To conclude, the use of animals for ART-research can be justified by the principles of subsidiarity and proportionality. In regard to subsidiarity, the use of animals remains necessary for specific forms of ART research. In regard to proportionality, taking account of different views of the moral importance of helping people to have a genetically related child and the importance of efficacy and safety research as part of responsible innovation, the limited infringement of animal wellbeing involved in ART research is outweighed by the benefits. Yet, this justification should always be assessed on a case-by-case basis, as the proportionality of animal research in practice still depends on the specifics of the case. As animal research inevitably comes at a moral price, it remains important to constantly consider to what extent the ideals of the 3Rs can be better achieved. Despite the fact that a full replacement of animal research on new ARTs is not realistic, possible alternatives including the use of human embryos for research should be considered.

Chapter three

In the third chapter, I questioned whether there is an ethically preferred order of preclinical research steps on new ARTs. Responsible innovation in MAR requires preclinical safety research, including the use of animal models and human embryos. However, the moral sensitivity of human embryo research has led to regulations and guidance requiring that human embryos may only be used for research that cannot also be conducted with animals. I call this the ‘use animals first’ (UAF) rule. In the field of ART research, this translates into the notion of an ‘ideal chain’ of consecutive preclinical research steps, where research using human embryos may only be considered as a further step after promising results have first been obtained in animals, or when it is clear that animal studies cannot contribute even in the slightest sense to answering the question.

From a purely scientific perspective, there is no such thing as a universally preferred order of research-steps. Scientifically speaking, the best pathway in research is to always choose the optimal way of answering a specific research question. The UAF-rule is not necessarily at odds with this. For research questions where the use of human embryos is not suitable to start with, it is clear that research should start in animal models. Also in cases where human embryos are suitable and the same research with animal embryos would not be informative, there is no conflict. In such a case, the same question cannot be answered in research using animal embryos and thereby meets the subsidiarity principle, which means the UAF-rule does not stand in the way of directly moving to human embryo research. In practice, however, the UAF-rule is interpreted more extensively, resulting in the notion of the ‘ideal chain’. This becomes problematic when research ethics committees demand that animal research should precede the use of human embryos, even in cases where that prior step is expected to be only slightly relevant to answer the research question. As a result, scientists are withheld from choosing which pathway is scientifically desirable on a case-by-case basis. This is, however, ethically problematic in so far as it may lead to a waste of valuable research time and money, but also because the benefits to be gained through that research may not outweigh the inevitable adverse impact on animal wellbeing, thus violating a core principle of the ethics of animal research.

Given these scientific and ethical drawbacks of the UAF-rule, what then are the possible moral arguments for upholding it? A first possible moral argument is that human

embryos would 'stand to lose more' than animals, when used for ART research, and should therefore be the least favored research material. To assess this argument, I investigated what animals stand to lose when being used as research material and how this compares to the fate of human embryos. The strong consensus acknowledges that, in contrast to mere tissues and cells, animals have a moral standing. This means that we are obliged to consider their needs and interests when making decisions affecting them (Warren, 1997). Philosopher Jeff McMahan describes this as the 'morality of interests' (MoI), which refers to how the interests of beings at the receiving end of our actions constrain what we may do to them (McMahan, 2002). In animal research on new ARTs, all animals suffer some extent of discomfort and/or pain. How does this compare to the use of human embryos for research? As preimplantation human embryos are not capable of experiencing pain or psychological distress, embryo research does not have an impact on their wellbeing. In addition to potential suffering, premature death can also harm interests (at least in humans) and taking a life is generally considered to be the greatest possible infringement on the victim's interests. This infringement depends on the level of losing a psychologically complex life and the level of connection to the future life, dependent on psychological features. In this context, it seems less wrong to kill animals such as mice, than animals such as non-human primates. In regard to human embryos, such deprivation would require a level of connection of a present self to a future self that goes beyond what can be attributed to an embryo. Therefore, the UAF-rule cannot be supported by moral arguments based on MoI.

A second possible line of argument for the UAF-rule is that, in contrast to animals, a special moral status can be attributed to human embryos regardless of whether they can be said to have interests. McMahan describes this as 'Morality of Respect' (MoR), which concerns what we owe to persons. As animals generally fall below the 'person threshold' (DeGrazia, 1997), they are not entitled to the same degree of protection as human individuals. This means that, whereas the use of persons as 'mere means' is morally unacceptable, this does not hold for beings below this threshold. With most animals falling below this threshold (some consider great apes 'borderline persons'), their use for entirely instrumental purposes, such as most ART research, can be acceptable under conditions.

What does this mean for the use of human embryos for research? However, most people accord a more limited moral status to the human embryo, which is not enough to grant

them the full protection of the MoR. The widely supported view is that the special moral status of human embryos and fetuses is initially quite low in comparison with that of persons, and that this status only increases with *in utero* development. Still, this relatively low status is in comparison with persons, not with animals, to whom such a special moral status is not ascribed at all (apart perhaps from great apes as borderline persons (DeGrazia, 1997)). Those defending the UAF-rule, may do so based on this presumed difference between human embryos and animals.

A third possible argument supporting the UAF-rule concerns oocyte donation, at least in cases where the use of spare embryos is not an option. Oocyte donation for the benefit of science is a morally sensitive practice, as it poses possible risks on donors without a direct benefit for themselves. For example, in order to produce mature oocytes, the donor must undergo hormone stimulation and oocyte pick-up. The procedure is burdensome and may entail physical risks, including a risk of the woman developing ovarian hyper stimulation syndrome (OHSS). In addition to physical risks, there are also possible social risks. For example, financial compensation schemes or so-called egg-sharing procedures (in which IVF-patients are offered free or reduced-cost treatment in return for donating part of the mature oocytes obtained in the process (Blyth, 2002)) have been criticized for inviting exploitation of vulnerable women. Yet, with careful protocols, physical risks can almost completely be avoided (Devroey et. al., 2011). Likewise, social risks can be minimized with carefully designed protocols and practices of information, counseling and consent (Mertes & Pennings, 2006). Ideally, oocytes that are obtained in the context of a fertility treatment but that cannot be used for reproductive purposes (due to failed fertilization, abnormalities or insufficient maturation, for example) are used for research, thus avoiding ovarian stimulation with the sole purpose of donation for research. Certainly, the challenges of responsibly obtaining mature human oocytes for preclinical embryo research need to be taken seriously and are therefore an argument for using spare embryos whenever the research question can be answered with these as well. However, given the possibility to avoid physical risks and minimize social risks, it is difficult to maintain that these challenges are so large as to settle the issue in favor of the UAF-rule.

In conclusion, it appears that the UAF-rule is based on the view that human embryos have a special moral status, without including the MoI perspective. When both perspectives

are included, it is less clear that the UAF-rule is morally tenable. As no overriding moral arguments supporting the UAF-rule were found, I recommend that in legal documents and professional guidance, the UAF-rule is abandoned and replaced by a justificatory framework that does not a priori favor animal studies over the use of human embryos in research. The notion of an ideal chain of ART research steps should be modified accordingly. A core message must remain that the step to the clinic can only responsibly be made after adequate preclinical safety studies. As the ill-founded notion of the greater moral sensitivity of human embryo research may stand in the way of this, it should be abandoned to avoid waste and increase the relevance of preclinical safety research. In line with this, research ethics committees assessing protocols for the use of human embryos in preclinical ART research should refrain from requiring prior animal studies that are not expected to be (sufficiently) informative or that imply a greater moral cost than embryo research.

Chapter four

The field of MAR has been criticized for introducing new ARTs without systematic preclinical efficacy and safety studies and without follow-up of children conceived with those technologies, thereby posing risks on patients and offspring (Dondorp & de Wert, 2011; Harper et. al., 2012, 2017; Sharpe, 2018). In the fourth chapter, I assessed whether this critique is justified and if so, how the field of MAR should then deal with offspring safety risks of introducing new ARTs. To answer the first part of this question, I assessed whether offspring safety considerations did or did not play a role in the history of introducing new ARTs. I examined three case-studies for this: two case-studies from the past, intracytoplasmic sperm injection (ICSI) and preimplantation genetic testing (PGT), and one contemporary, mitochondrial replacement therapy (MRT).

ICSI is an *in vitro* fertilization (IVF) procedure in which a single sperm cell is injected directly into the cytoplasm of an egg and was introduced in the 90s, primarily for the treatment of severe cases of male subfertility. Although, at the time, there was some awareness of uncertain offspring safety risks, hardly any preclinical research was performed. Yet, some safety measures were taken once ICSI was introduced, as in many fertility centers, the use of regular ICSI was subjected to (more or less strict) conditions (e.g. not using non-motile sperm). Only in the Netherlands, two surgical variants of ICSI, microsurgical

epididymal sperm aspiration (MESA) and testicular sperm extraction (TESE), developed for more severe cases of male subfertility, were (temporarily) prohibited due to lack of safety evidence. Although follow-up is not a safety measure in the strict sense, it may ultimately lead to improving the technique and decrease safety risks for (multigenerational) offspring. In the case of ICSI, only few centers conducted follow-up research, including UZ Brussel.

PGT enables the selection of embryos based on an evaluation of their genetic composition. Three types of PGT can be distinguished. PGT for monogenetic disorders (PGT-M), and PGT for structural rearrangements (PGT-SR) are indication-based applications aimed at helping at high risk couples to prevent transmitting a serious genetic disease to their offspring. These two types were introduced in the late 1980s and are historically referred to as preimplantation genetic diagnosis (PGD) (Handyside, Kontogianni, Hardy, & Winston, 1990). The third type, PGT for aneuploidy (PGT-A), was introduced to screen the chromosomal constitution of IVF-embryos with the aim of increasing the chance of successful pregnancy. This type was historically called preimplantation genetic screening (PGS) and was introduced in 1995 (Verlinsky et. al., 1995). Before introduction, concerns about offspring safety risks were widely shared (e.g. increased pregnancy loss). Although some preclinical research was performed, information on safety was limited when the technique was introduced into clinical practice (De Vos & Van Steirteghem, 2001). Furthermore, few safety measures were taken in clinical application and only few centers conducted follow-up. Additionally, PGT-A is increasingly offered beyond its indication area, without evidence of effectiveness (Mastenbroek & Repping, 2014; Penzias et. al., 2018).

MRT is a form of IVF in which the future child's mitochondrial DNA comes from a donor, and is aimed at helping women prevent transmission of mitochondrial DNA defects to their offspring. It is a relatively new technique and is currently only officially allowed in the UK. For MRT, there was a high level of risk awareness and extensive preclinical safety and efficacy research has been performed. The introduction of MRT varies widely. In some countries, MRT is prohibited (e.g. US/FDA), while in other countries, there is no MRT legislation (Ukraine). In the UK, MRT is allowed under strict conditions. Examples of this are extensive safety measures in clinical study (e.g. only offered to patients with a high risk of passing a serious mitochondrial disease) and follow-up research.

To assess whether and where exactly the field may have fallen short, I used Per Sandin's four dimensions relevant to dealing with uncertain risks. These dimensions are presented in the context of his discussion of the so called 'precautionary principle'. Following Sandin, the general structure of the precautionary principle can be reproduced as: "If there is (1) a threat, which is (2) uncertain, then (3) some kind of action (4) is mandatory" (Sandin, 1999).

The case studies show that attention has indeed been paid to both the threat and uncertainty dimensions. Concerns about the safety of new techniques were on the agenda of debate and sometimes led to proposals for concrete safety measures (e.g. parameters for sperm selection at ICSI). Yet, the case studies also show that, in light of the uncertainty of the threat (both in terms of chance of actualization and of seriousness), the step to the action and demand dimensions often did not take place.

A spectrum of how to deal with offspring safety risks can be identified. On one end of the spectrum lies the strong interpretation of the precautionary principle. In this context, a certain technique should not be introduced as long as there is uncertainty regarding offspring safety risks. Although not explicitly mentioned, this reasoning seems to underlie the Dutch moratorium on MESA and TESE. The ban on these variants of ICSI was only lifted in 2012 (MESA) and 2014 (TESE), after reassuring safety data became available. Here, the action dimension and the command dimension can be identified. Likewise, both dimensions can be identified in the current ban on the reproductive use of MRT by the Food and Drug Agency (FDA) in the United States. Yet, both these cases seem to be the exception in our studies.

On the other end of the spectrum of dealing with risks lies the 'innovation principle'. Although the four dimensions are originally presented in the context of the precautionary principle, using them for my analysis does not require committing to the precautionary principle as a framework for introducing new ARTs per se. The so called 'Innovation Principle', where "the call is for the benefits of innovation to be weighed against known harm" (Read & O'Riordan, 2017), also deals with uncertain risks. However, in contrast to the precautionary principle where uncertainty should lead to action, the innovation principle insists that only *known* harm can be a ground for limiting the application of new technologies. This approach is founded on the view that unhindered innovation is of great importance, and of which the benefits would be frustrated by unnecessary measures of precaution. Yet, this

approach may invite the view that it is acceptable to introduce new techniques on a basis of trial-and-error, which also seems to underlie the forms of ‘indication creep’ identified in the case studies. Using ICSI for non-male infertility problems, PGT-A to increase pregnancy chances, and MRT in the context of ‘egg-rejuvenation’ are examples of this.

If the innovation principle stands for giving innovation a green light in the face of uncertainty, and strong versions of the precautionary principle would stand for giving it a red light, intermediate, or weaker, forms of the precautionary principle may be understood as inviting an amber light policy, allowing innovation to ‘proceed with caution’. This means that although uncertainty can be considered a valid reason for taking measures, these measures do not have to be prohibitive. What measures need to be taken should be assessed based on proportionality: the expected advantages should be weighed against the uncertain risks. Both safety and efficacy should be part of this assessment, as safety risks are always disproportional when there is no clear benefit for patients and offspring.

To conclude, although offspring safety concerns about introducing the relevant technologies have been on the agenda of the debate in each of the cases, they also show a lack of systematic safety measures to deal responsibly with uncertain risks. Relevant recommendations from ESHRE (action dimension) seem to be understood as non-committal suggestions rather than as an articulation of a shared professional responsibility (action dimension). While this may not be a problem from the commercial perspective of a large part of the reproductive field, the innovation principle-approach seems at odds with the notion that professionals in assisted reproduction have a responsibility to take accountability for the welfare of the children they bring into being. Although strong (prohibitive) forms of the precautionary principle seem difficult to justify given the equally morally important interests of those dependent on medical help for having children, I suggest a policy of proceeding with systematic caution. This means that safety risks associated with introducing new ARTs should be minimized as much as possible. ESHRE’s framework considering responsible innovation on MAR can be used as an example for this. Moreover, governments should play an important role by formulating legal frameworks in which the professional group can implement its responsibility in a self-regulated way.

Chapter five

According to ESHRE's normative framework on responsible innovation, follow up (ideally long term) of children conceived through medically assisted reproduction (MAR) should be an integral part of introducing new assisted reproductive technologies (ARTs). Still, very few centers actually conduct it. With potentially risky new ARTs on the horizon, these recommendations need to be taken more seriously. Without the illusion of being exhaustive, I explored two of the main ethical considerations related to responsibly conducting follow-up studies on ARTs: the proportionality of the inevitable burdens and risks for families participating, and the implications of the principle of respect for autonomy.

Follow-up research in the case of MAR concerns the monitoring of both the patients' and offspring's health over time after treatment. In this thesis, I solely focused on follow-up research in offspring, as this is unique about MAR follow-up. Follow-up studies (both short and long-term) on children conceived after assisted reproduction may be important for four different groups of individual stakeholders, as well as for society as a whole. First, follow-up may serve as a safety measure to decrease possible health risks for offspring born after the researched ART. As introducing new reproductive treatments always involves uncertain risks for children thus conceived, follow-up may provide early detection of possible risks for those children. Second, when long-term follow up also concerns examining the health of the descendants of MAR-children, this generation also benefits from possible early detection. In both cases, the advantage of early detection is dependent on whether the findings are actually actionable, meaning that treatment is possible. Third, if findings are used to make MAR-procedures safer, this benefits future children born after ART by reducing the risk of being born or having to live with ART-related malformations or disorders. Fourth, follow-up contributes to informed decision making for reproductive patients, as it causes a better understanding of the possible harms and benefits of a certain reproductive technique. For society at large, follow-up may contribute to safeguarding public health. As nowadays, children born after assisted reproduction are a substantial part of the population, even small health risks as a result of ART could have a large impact on public health.

Although some centers have contributed to follow-up studies extensively, especially at UZ Brussel, these centers are still the exception. At the same time, in many cases where follow-up was actually conducted, studies had considerable methodological limitations

(Mulder et al., 2018; Sutcliffe & Ludwig, 2007). Practical difficulties, such as low turnout rates and scarcity of financial resources, are often mentioned as a reason behind the lack of (methodologically strong) follow-up research worldwide.

The first ethical issue related to responsibly conducting ART follow-up I am discussing concerns proportionality. I questioned whether the possible benefits to be obtained from the relevant study clearly outweigh the possible risks and burdens for the participants. ART follow-up often entails both medical and psychological tests and are recurrent. Participating in such follow-up tests may lead to medical risks and psychological burdens. As most tests concern non-invasive check-ups, medical risks are usually minimal. In some cases, however, tests may be more invasive (e.g. blood drawing, tissue typing requiring biopsy). Psychological burdens may include, but are not limited to, psychological discomfort related to possible examination of reproductive organs and questions related to sexual development and behavior, feeling different from other children, and the potential of discovering serious health risks. These medical risks and psychological burdens are particularly troublesome because MAR follow-up mainly entails research with minors. Research with minors is accepted when there is potential benefit for themselves, and risks are minimized and outweighed by this benefit (Doek et al., 2009; Council of Europe, 1997; Council for International Organizations of Medical Sciences, 2017). If this individual benefit is not evident, research with minors is only allowed if 1) the research question can only be answered by conducting research with minors (criterion of group-relatedness) and 2) the burdens and risks for those minors are minimal. Because only research with children can give insight into the possible impact of being MAR-conceived on children's development, MAR follow-up meets the criterion of group-relatedness. Yet, the level of risk and burdens, and thereby the acceptability of follow-up in minors, depends the specific research question and the specific research phase. This leads to a difficult tension: while the need for comprehensive safety data is greatest at the experimental stages of a potentially risky new technology, the possibility to collect such data through follow-up studies in children is limited by the requirement of minimal risks and burdens. In established treatment, evidence on safety is already quite extensive. Although risks associated with the researched technique are considered acceptable by then, which also makes follow-up studies with children more acceptable, safety studies are less crucial at that point than at the experimental stages.

If the proportionality principle can be met, a further criterion in the ethical examination of ART follow-up concerns the implications of the principle of respect for autonomy for ART follow-up research. The decision to participate in research should not only be informed, but should also be unconstrained and voluntary (Faden & Beauchamp, 1986). In practice, this means that children who are mature enough to have an adequate understanding of the nature of follow-up research and its implications should be allowed to choose whether or not to participate in such studies. At the same time, children who are not yet competent in this sense, should not be burdened with choices that are beyond their capacity for autonomous agency (Hein & Jöesaar, 2015). In such cases, parents should decide for them.

What does the condition that follow-up should be voluntary mean for the possibility to actively encourage parents to give consent for their children participating in follow-up? This depends on whether MAR-families (parents and children) can be said to have a certain responsibility to participate in follow-up. If so, it can be argued that encouraging them may be acceptable under conditions.

There are two possible grounds for ascribing such a responsibility to MAR-families. The first ground concerns parental responsibility. If a new ART is still experimental or innovative, parents have a *prima facie* responsibility towards their children to enroll them in follow-up research that may lead to early detection of health problems or health risks that allow for timely treatment or prevention (Ishii, 2019). With 'prima facie', I mean that the research should still be proportional, taking account of risks and burdens. In such a case, reminding parents and their family of this responsibly is not at odds with their autonomy, as long as they are still free to decline. If a new ART is introduced in the context of formal research, professionals may require an agreement to participate in follow-up as a pre-treatment condition. Still, this agreement can never be enforced, because research participants always have the right to withdraw from a research project. When new ARTs become established treatment, this means there is sufficient evidence that there are no serious concerns about its immediate safety. Using 'parental responsibility' as a ground for ascribing a moral responsibility to participate in MAR follow-up would then lose most of its force. A second possible ground is reciprocity. According to bioethicist John Harris, we have a moral duty to participate in medical research because of two reasons (Harris, 2005). First,

by volunteering as research subjects, we can save others from serious harm without excessive cost to ourselves. Secondly, all members of modern Western societies benefit from advances in healthcare and they therefore have a moral responsibility to contribute to the social practice which produces these benefits. These arguments can be criticized, however. As to the first argument, it can be argued that the connection between participating in research and possibly saving human lives as a result is far too indirect to serve as an argument for a moral responsibility. Regarding the second argument, although it may be fair to require some form of reciprocation when benefiting from medical advances, it does not specify what this should mean in practice. In light of this, it has been argued that everyone has an 'imperfect moral duty' to further the general good and that participating in scientific research is just one way to do so (Shapshay & Pimple, 2007). In the case of follow-up research, specifically, it may be an additional argument that only those who have undergone the researched procedure can be asked to participate. Being better placed than most others to make this specific contribution to the general good may be a moral reason for at least seriously considering such a request.

To conclude, as adequate follow-up is essential for responsible innovation in MAR, the reproductive field should put more effort in actively implementing ESHRE's normative framework considering the execution of MAR follow-up. Professionals may encourage MAR families to participate in follow-up research if that study meets the condition of proportionality. In the context of formal research, pre-treatment agreement to participate in follow-up may be presented as a (non-enforceable) condition for access. In case of established treatment, still the fact that MAR parents and children are better placed than others to contribute to the general good may be a legitimate ground for encouraging them to participate.

Focus group with fertility patients

As discussed in the Introduction of this thesis, I aimed to conduct several focus group interviews to investigate the attitudes of fertility patients towards potentially risky innovations in MAR. Despite extensive recruitment efforts, only one focus group interview could be performed with the limited number of eight participants. Due to this limited turnout rate, and due to time constraints, results have not (yet) been published. Yet, I believe one of

the preliminary conclusions from the focus group interview nevertheless deserves specific attention, because of its relevance to the field of MAR.

One of the key questions that was asked to the participants explored what level of risk for ART conceived offspring they considered acceptable. The study of Dancet et. al. shows that fertility patients are more accepting towards preclinical research (e.g. embryo and non-human primate research) on new ARTs than the general public (Dancet et. al., 2011). One explanation of this acceptance is the fact that they are potential beneficiaries of the results of the preclinical research. I therefore expected that fertility patients may also be more accepting of risks for themselves and their future offspring, because of their strong wish for a child. This means that what they consider to be an acceptable risk in MAR might be different to what non fertility patients would consider to be an acceptable risk.

The outcome of the focus group was ambivalent. The expectation that fertility patients would be more accepting to take risks for themselves and future offspring was partially supported by the results of the focus group. Participants mentioned they were indeed willing to take risks they considered relatively high. The reasoning that participants mentioned behind this willingness to take higher risks was two-fold. First, they mentioned their strong desire for having a (genetically related) child. Secondly, their boundaries shifted the further they were involved in the treatment trajectory. Surprisingly, however, participants themselves were very aware of their own bias and acknowledged multiple times that they were too emotionally involved to make unilateral decisions in fertility treatment. Interestingly, participants mentioned that the level of risk they were willing to take was not necessarily a level of risk that society as a whole should accept. For example, a few participants stated that they themselves were willing to take higher risks, but considered it irresponsible to set this level of risk as a universal benchmark. They believed that accepting this level of risk could endanger public health. Yet, participants also underlined the importance of including individual wishes and boundaries, as no two fertility patients are the same. In line of this, shared decision-making was considered of great importance.

As patients acknowledged their own bias considering risk-taking, participants proposed a multi-stakeholder approach to responsibly introduce new reproductive techniques. First, the government should provide legal frameworks through legislation and regulation. Researchers are perceived as co-responsible for safely introducing new

reproductive techniques, within these legal boundaries. Second, healthcare professionals should provide information and recommendations to potential fertility patients, because they have medical knowledge. Last, couples should be able to determine whether a technique is suitable for them or not in consultation with a healthcare professional.

2. Wider implications and agenda setting

In this part, the wider implications of the thesis are discussed. Furthermore, themes for further discussion are addressed. The four subsections regard: animal research, human embryo research, practical implications for the normative framework on responsible innovation in MAR and the question ‘how safe is safe enough?’.

Animal research

Animal research is, and may always be, necessary for responsible innovation in MAR, aiming at risk reduction for children conceived after ART. At least for the time being, there are no adequate alternatives to replace all animal research on ARTs. Taking account of different views of the moral importance of helping people to have a genetically related child and the importance of safety research as part of responsible innovation, the use of animals for ART efficacy and safety research is, in principle, justified. In principle, because the proportionality of concrete instances of animal research depends on the specifics of the case.

Many argue that the type of animal used for research is one of the specifics affecting the subsidiarity and proportionality criterion in animal research. Often, a distinction is made between the research use of non-human primates and other animals. The 2010 EU Directive on the protection of animals used for scientific purposes provides a clear example of this. The Directive states that; “due to their genetic proximity to human beings and to their highly developed social skills, the use of non-human primates in scientific procedures raises specific ethical and practical problems in terms of meeting their behavioral, environmental and social needs in a laboratory environment. Furthermore, the use of non-human primates is of the greatest concern to the public” (European Commission, 2010, p. 17). Apparently, the use of non-human primates for research needs additional justification.

Importantly, when discussing the research use of non-human primates, the Directive makes a further distinction between the use of great apes, like gorillas, chimpanzees and orangutans, on the one hand, and the use of other non-human primates on the other. Whereas the research use of the latter is only allowed with additional justification, research with great apes is even further restricted. According to the Directive, this is only allowed for research that is “aimed at the preservation of those species and where action in relation to a life-threatening, debilitating condition endangering human beings is warranted, and no other species or alternative method would suffice in order to achieve the aims of the procedure” (European Commission, 2010). This clearly leaves no room for using great apes for ART-research aimed at benefitting humans. Moreover, in the literature the case is made for completely ruling out the instrumental use of great apes for research, based on the argument that as ‘borderline persons’ (DeGrazia, 1997), they should be regarded as on the same side of the ‘personhood threshold’ as humans, and may therefore not be used as mere means. Siding with the strong consensus in this matter, I will not consider the use of great apes for ART research. In the below argument my use of the term ‘non-human primates’ should therefore be understood in the restricted sense of excluding not just humans, but great apes as well.

This leaves us with the distinction between non-human primates and other animals as morally different categories of research animals. According to the Directive, non-human primates may only be used if it can be shown that the aim of the research procedure cannot be realized with animals belonging to other species (European Commission, 2010). As this seems to introduce a distinction between ‘higher’ and ‘lower’ animals, the question is how such a distinction can be accounted for without ending up with a traditional anthropocentric tree-of-life image, where a gradualist understanding of moral status is determined by closeness to humans as the crown of creation.

One possible approach is to say that all sentient life below the personhood threshold has the same moral status, reflecting the fact that having interests makes a being a possible moral patient of human activity. Following this line of reasoning, morality requires us to consider the needs and interests of animals whenever making decisions affecting them, but without implying a distinction between lower and higher animals in a moral sense. Although it is still possible to use those terms, they rather refer to a lesser or greater complexity of the mental and social lives of these animals, implying that any further distinction between, say,

rodents and non-human primates as potential research models is better understood as referring to the differential degree of infringement of their wellbeing that using them for research would entail. The more the use of specific animals in research leads to pain, discomfort, a negative psychological impact, or an infringement of species-specific behavior, the greater the human benefits must be in order for the research in question to be considered proportional. The dichotomy in the Directive may thus be understood as pointing to the fact that the higher level of psychological and social complexity exhibited by non-human primates raises a higher moral barrier to using them for research purposes than is the case with other animals. However, as the criterion concerns impact on species-specific wellbeing, it would seem to apply beyond the dichotomy as well. Using the same criterion, one might say that for similar reasons of differential complexity, research using, say, pigs is morally more challenging than doing the same research in mice. Others may reason that precisely this differentiation between animals with lives of lesser and greater complexity entails a hierarchy in moral terms as well. This would then translate in a differential ascription of moral status between animals of different species, with pigs, say, having a higher moral status than sheep.

Whichever of these lines of reasoning is followed, it does not necessarily mean that animal research should preferentially be done with animals having simpler species-specific needs (or, for that matter, a lower moral status). For specific research questions, the use of animals with more complex species-specific needs may have additional value regarding efficacy and safety information that cannot be achieved with other animal models. Given the responsibility of MAR-professionals towards their patients as well as to the children they help bring into the world, the ethical justification for the use of animals with more complex needs is highly dependent on the scientific value of the relevant research. If the use of animals with more complex needs yields significantly more robust data on safety and efficacy, the use of such animals may be more proportional. Using animals with simpler needs for research when it is known that these models do not yield robust data should even be regarded as ethically problematic in so far as this violates a core principle of the ethics of animal research, namely that the inevitable adverse impact on animal wellbeing should be proportional to the importance of the benefits to be gained through that research (Jans et. al., 2018). Therefore, we should scrutinize current presumptions on the proportionality of using specific animals

for research and be open for change, recommending, on conditions, also the use of animals with more complex needs, such as non-human primates, if this would offer a considerable benefit.

Human embryo research

Currently, in the Netherlands, there is an extensive debate on the acceptability of creating human embryos for research, which is categorically banned by the Dutch Embryos Act (CBO, 2002). It is permissible to use spare embryos for research, under conditions. As stated in article 10, embryo research is only acceptable when it is likely that the research will lead to new insights in medical science. Furthermore, research using embryos is only permissible when it cannot be performed using other methods of scientific research (CBO, 2002). Scientists and ethicists have been arguing for years to amend the Embryos Act in order to also (conditionally, of course) allow the creation of embryos specifically for research. The reports of both the first and the second Evaluations of the Embryos Act recommended that the categorical ban on creating embryos for research should be lifted, as it blocks important research (Olsthoorn-Heim et. al., 2006). Given the political situation, taking account of the great influence of the Christian parties in the relevant Dutch coalitions, the ban was, however, not lifted. In 2016, the Minister of Health announced her intention to amend the Act in order to allow the creation of embryos specifically for preclinical safety studies regarding particular reproductive technologies (Schippers, 2016). Presently, the current coalition of two liberal parties (VVD and D66) and two Christian parties (CDA and ChristenUnie) agreed to not make any radical medical ethical decisions, which means that a liberalization of the Embryos Act is not likely to happen soon. Yet, the government decided to organize five public consultations to investigate the public's attitude concerning the Dutch Embryo Act, one of them regarding the acceptability of specifically creating embryos for research. A recent study concerning the attitudes of Dutch citizens on embryo research shows that a slight majority is accepting towards the creation of embryos for research. Arguments concerning oocyte donation and the desirability of new reproductive techniques weighed heavily. Furthermore, the study points out that a broader public debate is necessary to better include the variety of perspectives in future political decision making (Gouman et. al., 2020).

As discussed earlier in this thesis (chapter 3), the prohibition of creating human embryos for research hampers responsible innovation in MAR. As spare embryos are already a couple of days old, it is impossible to use them for research on the early development of embryos. Moreover, to study the safety and efficacy of new reproductive techniques, for example oocyte vitrification, embryos created with those specific techniques need to be investigated. For both research purposes, the use of embryos specifically created for research is necessary.

What are the possible moral arguments behind the prohibition of creating embryos for research and are they valid? The moral debate on the creation of embryos for research focuses on whether a significant moral difference can be made between using spare embryos for research and creating research embryos. If there is no moral difference between using spare and creating research embryos, it could be argued that the creation of research embryos should be allowed just like the use of spare embryos is. This issue has been a longstanding topic of debate in the field of embryo research, whereby one position used the distinction to defend the prohibition of research embryos (van Beers, 2016), while others argue that such a moral distinction is either indecisive or simply invalid (te Braake, 1998; Devolder, 2005, 2013; Gezondheidsraad, 1998; de Wert & Geraedts, 1988).

The authors who point out a difference between using spare embryos and specially created embryos for research do not aim at a difference in moral status considering internal factors (i.e. factors inherently connected to the embryo), but argue that a moral difference can be made between the *intention* of creating spare embryos and creating research embryos (van Beers, 2009). For instance, van Beers has argued that the creation of embryos for research introduces a new category of human life that is only intended for instrumental use instead of serving a 'relational value' (i.e. reproductive purposes), which in her view is at odds with the protection that embryos deserve (van Beers, 2016). The argument here is that creating research embryos instrumentalizes these embryos, meaning they are reduced to a mere means from the very beginning. This would be at odds with the dignity of human life. In contrast, spare embryos were at least created with a relational value in mind, namely helping a couple with reproduction (van Beers, 2016). This argument is, however, incoherent. As argued by bioethicist Katrien Devolder, in standard IVF practices, embryos are just as much instrumentalized as when creating research embryos (Devolder, 2005, 2013).

Namely, in standard IVF, more embryos are created than can responsibly be transferred to the womb in the same cycle. The remaining embryos are frozen. If an attempt to achieve a pregnancy fails, women do not have to undergo the hormone therapy and egg retrieval procedure, because one or two of the frozen embryos can be thawed and used, thereby reducing health risks and discomfort. The argument of instrumental use can thus not be used to defend a moral difference between spare embryos and research embryos.

A response on the argument that, in standard IVF practice, spare embryos are also instrumentalized, is given by biomedical philosopher Timothy Murphy. He argues that embryo loss in fertility medicine using IVF is not intended, but only a foreseen side effect of the procedure, which is referred to as the principle of double effect. This differs from the intentional destruction of embryos that occurs in research (Murphy, 2012). Following the same line of argument, moral philosopher William FitzPatrick has argued that the moral difference between using spare embryos and creating embryos for research lies in a distinction between intention and foreseeability (FitzPatrick, 2003). According to FitzPatrick, IVF clinicians foresee that probably not all embryos will result in a child. However, they do not create embryos with the intention to destroy them. On the contrary, when creating embryos for research, we are “creating embryos with such an intention, exhibiting the distinctively exploitative and opportunistic attitude bound up with such a practice” (FitzPatrick, 2003, p. 31). These arguments, based on the principle of double-effect, are, however, invalid. As some ethicists have argued, the creation of spare embryos in woman-friendly IVF (in order to increase pregnancy chances, more embryos are created than can responsibly be transferred in the same cycle) cannot be understood as merely foreseen, because one could very well limit the creation of embryos to the amount (one or two) that can be responsibly transferred to the womb in one IVF-cycle (embryo-friendly IVF) (Devolder, 2013; de Wert, 1987). In a society that accepts woman-friendly IVF to be the standard IVF practice, double-effect reasoning cannot support the view that creating embryos for research instrumentalizes embryos and that creating more embryos than that will be used (spare embryos) does not. As to the intentional destruction of embryos, since a significant part of IVF-embryos will never be used for reproductive purposes, it cannot be argued that these embryos are not created with the knowledge that they will be destroyed. The fact that only part of IVF-embryos will eventually be transferred and have the chance to

result in a child, does not mean that not all embryos are instrumentalized from the beginning, because all embryos are created with the intention of facilitating the possibility to select the best of the batch for implantation.

Also arguments concerning risks for women donating have been given as a reason for making a moral distinction between using spare embryos and research embryos. It is argued that in egg cell donation, women are exposed to medical and social problems; medical, because of the hormone therapy and invasive IVF procedure without knowing the risks, and social, because of a potential exploitation of women potentially participating in egg donation, without any physical, personal, or financial benefits (Baylis, 2000; Gerrand, 1993). This can be seen as a reason to make a moral distinction between the use of spare embryos and research embryos. However, although the concerns should be taken very seriously in the debate on research embryos, it has been argued that the risks for egg cell donors are not as big as often suggested, and can and should be minimized (Winter, 2012). As to medical risks, ovarian hyper stimulation syndrome has been associated with the emergency of pregnancy (Devroey et. al., 2011). This means that women donating egg cells outside a reproductive procedure are less likely to be at risk. As to social risks, several legal conditions for egg cell donation can prevent the possible exploitation of women, such as the condition that women are not allowed to be paid for their donated egg cells (Olsthoorn-Heim et. al., 2006). Furthermore, with the current abilities to freeze oocytes, it is expected that a part of them will remain unused by the woman for whom they were originally stored. These leftover oocytes may be directed to research purposes (Mertes et. al., 2012).

In conclusion, arguments in favor of a moral distinction between using spare embryos and producing embryos for research on e.g. MAR do not seem to be convincing. If the use of spare embryos is justified for research on new reproductive techniques, this then means that also the use of research embryos is justified in cases where spare embryos are not sufficiently available, are of lower quality, or are outright useless given the research question at hand. If the responsibility to safely introduce new reproductive techniques is considered important, it would not only be inconsistent, but, even immoral to uphold the categorical prohibition of creating embryos for MAR research. Clearly, assessing whether a specific study is eligible for using research embryos requires scrutiny on a case-by-case basis. The moral harm of (creating and) destructing embryos and the risks for egg cell donors should be weighed

against the envisioned benefits (principle of proportionality). Furthermore, no morally less invasive options (including spare-embryos) should be available (principle of subsidiarity). On the basis of stringent legal criteria, the practice in Belgium, one of the countries where creating human embryos for research is conditionally allowed, is that 8% of all embryos used in scientific research are specifically created for the purpose (Pennings et. al., 2017).

Some critics argue that the research use of human embryos is replaceable, for example, by using stem-cell derived ‘synthetic embryo-like structures’ (Fransen, 2017). They, apparently, argue either that such synthetic embryo-like structures cannot be considered embryos or that, if these can or should be considered embryos, these have a lower moral value than ‘regular’ embryos. This would mean that the research use of such embryo-like structures would be less morally sensitive than the research use of ‘regular’ embryos and maybe even that the latter would in fact not (or no longer) meet the subsidiarity principle. This view invites two comments. Firstly, the question as to what constructs or entities (could) qualify as embryos needs further scrutiny. It would be premature (and probably a good example of ‘jumping to conclusions’) to a priori disqualify all different types of embryo-like structures as non-embryos. Secondly, although using synthetic embryo-like structures is, with no doubt, valuable to address particular research questions, it is, however, not useful for research on new reproductive techniques; in order to assess their efficacy and safety, research should be performed on embryos specifically created with those exact techniques. The use of embryos specifically created for research on new reproductive techniques is therefore still crucial for responsible innovation in MAR.

Practical implications for the normative framework of ESHRE

This thesis gives reasons for at least three amendments (one refinement and two adjustments) of the present normative framework for responsible innovation in MAR as proposed by ESHRE.

The refinement to be considered regards the added value and complexity of *transgenerational* follow-up studies. Some health effects may only present themselves in next generations. For example, in germline gene editing, more subtle adverse effects may only manifest in the second or third generation. Therefore, the National Academy of Sciences Committee on Human Gene Editing has recommended intergenerational monitoring

(National Academies of Sciences, Engineering, and Medicine, 2017). At the same time, transgenerational research carries several practical and moral difficulties. Practical difficulties may include lack of funding, large dropout rates and the fact that transgenerational research is very time-consuming. Ethical issues may include concerns regarding autonomy and privacy. For example, it can be questioned “whether it is ethically permissible in some broader sense to monitor individuals on the basis of their connection through family lineage to the participant in a medical experiment conducted decades previously” (Cwik, 2017). Furthermore, studies will require researchers to have access to key medical data for entire families over several decades (Cwik, 2017). More research is needed on how to deal with the need for transgenerational follow-up research on the one hand, and the (increased) practical and ethical difficulties it raises on the other.

The two adjustments I would like to recommend are the following. First, the recommendation that human embryos may only be used for research that cannot also be conducted with animals (the UAF-rule) should be abandoned, so as to avoid waste and to ensure robust preclinical research. It should be replaced by guidelines that do not a priori favor animal studies over human embryos research. Early developmental (e.g. gene expression) research should be performed directly on human embryos as that would bypass the uncertainties that go with animal-based studies. This does, however, not mean that animal research is not valuable. Multigenerational animal studies should be performed to investigate potential health effects of new reproductive technologies at multiple stages of offspring development. With potentially risky reproductive technologies on the horizon (such as reproduction using stem-cell derived gametes or genetically modified gametes or embryos), it is important that optimal use can be made of either type of morally sensitive research material. In practice, this means that research ethics committees assessing protocols for the use of human embryos in preclinical ART research should refrain from requiring prior animal studies if those are not expected to be informative.

Second, self-regulation lacks commanding power and is thereby often not followed in practice (chapter 4). Legislation is necessary to make sure that MAR is performed in a responsible way. Without legislation backing up self-regulation, fertility centers have the room to operate ‘carte blanche’. This is particularly problematic for a field that is to a large extent driven by commercial interests, often offering treatments on demand without

evidence of efficacy and safety. Due to risks for both offspring conceived by means of ART and public health, (inter)national legislation is necessary to make sure regulation is followed and treatments are efficient and safe. A system such as in the UK, where the HFEA makes sure fertility clinics and research centers comply with the law, or the Netherlands where a (to some extent) similar licensing system applies, may be an example of how those laws can be enforced in IVF clinics.

Finally: 'how safe is safe enough?'

Last, despite the fact that there is much literature on what research steps should be taken to introduce new ARTs, the issue of acceptable risk in medically assisted reproduction (MAR) has been largely ignored. In the current context, 'acceptable risk' refers to the stage when risks associated with the introduction of a new ART are considered acceptable and do not withhold a responsible introduction of an ART into clinical practice. As all reproduction comes with risks, a zero-risk benchmark is impossible. The question arises what should, then, be the risk evaluation standard, the threshold of acceptable risk, in the context of introducing new ARTs. ESHRE sets this threshold at: "high risk of serious harm" (Pennings et. al., 2007). The original context, however, of the reasoning behind this threshold is the question of whether or not to treat individual applicants of MAR when the would-be parents are a risk factor for the welfare of the child. For example, IVF-professionals may want to refrain from treating a couple when they are aware of former child abuse by this couple. ESHRE mentions three possible standards in evaluating whether to provide medical assistance. On one side of the spectrum, the 'maximal welfare' standard demands that no medical assistance is given when it can be expected that the life conditions of the future child will not be optimal. On the other side of the spectrum, the 'minimal welfare' standard states that assistance in reproduction is acceptable as long as the child would not be better off if it was not born. This standard reflects the so-called non-identity problem, which states that an individual cannot be harmed by an action that is a condition of their own existence, when that existence is (at least minimally) worthwhile (Parfit, 1987). The intermediate standard is 'reasonable welfare' and states that assistance is acceptable when it can be expected that the future person has all the opportunities to live a valuable life. ESHRE recommends the latter standard as threshold, which means that assistance is only acceptable when it can be expected that the

wellbeing of the child will not fall below the standard of reasonable welfare, and there is thus no expectation of high risk of serious harm. Yet, it can be questioned whether this threshold for acceptable risk also applies in the context of introducing new reproductive techniques, such as the reproductive use of stem-cell derived gametes or germline genome editing. Is this standard not too low? How about techniques entailing a high risk of a somewhat less than serious harm or moderate risks of serious harm? As cumulative risks of reproductive techniques may have adverse effects on multiple generations and adversely affect public health, a more stringent, less permissive, evaluation standard seems necessary. It should be investigated what then would be a responsible threshold of acceptable risk in assisted reproduction – ‘How safe is safe enough?’. More debate on this is necessary, in which multiple stakeholders, including professional societies and patient organizations, should have a role. As identified in the focus group with fertility patients, patients’ bias concerning risk taking should be taken into consideration, however, without excluding their perspective entirely. A multi-stakeholder approach to responsibly introduce new reproductive techniques, as proposed in the focus group, may be a balanced solution for this.

References

- Baylis, F. (2000). Our cells/ourselves: creating human embryos for stem cell research. *Women's Health Issues, 10*(3), 140-145.
- Blyth, E. (2002). Subsidized IVF: the development of 'egg sharing' in the UK. *Human Reproduction, 17*(12), 3254-3259.
- Brison, D. R., Roberts, S. A., & Kimber, S. J. (2013). How should we assess the safety of IVF technologies? *Reproductive Biomedicine Online, 27*(6), 710-721. doi:10.1016/j.rbmo.2013.09.006
- CBO. (2002). Dutch Embryos Act. *Alphen aan de Rijn*.
- Council for International Organizations of Medical Sciences. (2017). International ethical guidelines for health-related research involving humans. *International ethical guidelines for health-related research involving humans*.
- Council of Europe. (1997). *Convention for the protection of human rights and dignity of the human being with regard to the application of biology and medicine: convention on human rights and biomedicine; Oviedo, 4. 4. 1997: Conseil de l'Europe, Service de l'Édition et de la Documentation*
- Cwik, B. (2017). Designing ethical trials of germline gene editing. *New England Journal of Medicine, 377*(20), 1911-1913.
- Dancet, E. A., Van Empel, I., Rober, P., Nelen, W., Kremer, J., & d'Hooghe, T. (2011). Patient-centred infertility care: a qualitative study to listen to the patient's voice. *Human Reproduction, der022*.
- De Vos, A., & Van Steirteghem, A. (2001). Aspects of biopsy procedures prior to preimplantation genetic diagnosis. *Prenatal diagnosis, 21*(9), 767-780.
- De Wert, G., & Geraedts, J. (1988). IVF, pre-embryo-research en ethiek. *Metamedica, 67*, 106-123.
- DeGrazia, D. (1997). Great apes, dolphins, and the concept of personhood. *The Southern journal of philosophy, 35*(3), 301-320.
- DeGrazia, D. (2003). The Ethics of Animal Research: What are the Prospects for Agreement? In S. J. Armstrong & R. G. Botzler (Eds.), *The Animal Ethics Reader* (2nd ed., pp. 309-316). Abingdon: Routledge.

- Devolder, K. (2005). Human embryonic stem cell research: why the discarded-created-distinction cannot be based on the potentiality argument. *Bioethics*, 19(2), 1467.
- Devolder, K. (2013). Embryo deaths in reproduction and embryo research: a reply to Murphy's double effect argument. *Journal of Medical Ethics*, 39(8), 533-536. doi:10.1136/medethics-2012-101065
- Devroey, P., Polyzos, N. P., & Blockeel, C. (2011). An OHSS-Free Clinic by segmentation of IVF treatment. *Human Reproduction*, 26(10), 2593-2597.
- Doek, J., Breuker, K., Knibbe, C., Offringa, M., Passchier, J., & Steinkamp, N. (2009). Advies medisch-wetenschappelijk onderzoek met kinderen. *Den Haag*.
- Dondorp, W., & de Wert, G. (2011). Innovative reproductive technologies: risks and responsibilities. *Human Reproduction*, 26(7), 1604-1608. doi:10.1093/humrep/der112
- European Commission. 2010. Directive 2010/63/EU on the protection of animals used for scientific purposes. In 2010/63/EU, ed. European Commission. Ispra, Italy.
- Faden, R. R., & Beauchamp, T. L. (1986). *A history and theory of informed consent*: Oxford University Press.
- FitzPatrick, W. (2003). Surplus embryos, nonreproductive cloning, and the intend/foresee distinction. *Hastings Center Report*, 33(3), 29-36.
- Fransen, R. (2017, August 3, 2017). Embryo's zijn te repareren. *Nederlands Dagblad*. Retrieved from <https://www.nd.nl/cultuur/wetenschap/636810/embryo-s-zijn-te-repareren#push>
- Gerrand, N. (1993). Creating embryos for research. *Journal of applied philosophy*, 10(2), 175-187.
- Gezondheidsraad. (1998) IVF: afrondende advisering Rijswijk: Gezondheidsraad: Commissie Herziening Planningsbesluit IVF.
- Gouman, J., Vogelesang, S., & Verhoef, P. (2020). *Gewicht in de schaal – Nederlanders over onderzoek met embryo's*. Den Haag: Rathenau Instituut.
- Handyside, A. H., Kontogianni, E. H., Hardy, K., & Winston, R. M. (1990). Pregnancies from biopsied human preimplantation embryos sexed by Y-specific DNA amplification. *Nature*, 344(6268), 768.

- Harper J, Jackson E, Sermon K, Aitken RJ, Harbottle S, Mocanu E, Hardarson T, Mathur R, Viville S, Vail A. (2017). Adjuncts in the IVF laboratory: where is the evidence for 'add-on' interventions? *Human Reproduction*, 32(3): 488-491.
- Harper, J., Magli, M. C., Lundin, K., Barratt, C. L., & Brison, D. (2012). When and how should new technology be introduced into the IVF laboratory? *Human Reproduction*, 27(2), 303-313. doi:10.1093/humrep/der414
- Harris, J. (2005). Scientific research is a moral duty. *Journal of Medical Ethics*, 31(4), 242-248.
- Hein, V., & Jöesaar, H. (2015). How perceived autonomy support from adults and peer motivational climate are related with self-determined motivation among young athletes. *International Journal of Sport and Exercise Psychology*, 13(3), 193-204.
- Ishii, T. (2019). Should Long-Term Follow-up Post-Mitochondrial Replacement be Left up to Physicians, Parents, or Offspring? *The New Bioethics*, 25(4), 318-331.
- Jans, V., Dondorp, W., Goossens, E., Mertes, H., Pennings, G., & de Wert, G. (2018). Balancing animal welfare and assisted reproduction: ethics of preclinical animal research for testing new reproductive technologies. *Medicine, Health Care and Philosophy*, 1-9.
- Mastenbroek, S., & Repping, S. (2014). Preimplantation genetic screening: back to the future. *Human Reproduction*, 29(9), 1846-1850.
- McMahan, J. (2002). *The ethics of killing: Problems at the margins of life*: Oxford University Press on Demand.
- Mertes, H., & Pennings, G. (2006). Oocyte donation for stem cell research. *Human Reproduction*, 22(3), 629-634.
- Mertes, H., Pennings, G., Dondorp, W., & de Wert, G. (2012). Implications of oocyte cryostorage for the practice of oocyte donation. *Human Reproduction*, 27(10), 2886-2893.
- Mulder, C. L., Serrano, J. B., Catsburg, L. A., Roseboom, T. J., Repping, S., & van Pelt, A. M. (2018). A practical blueprint to systematically study life-long health consequences of novel medically assisted reproductive treatments. *Human Reproduction*, 33(5), 784-792.
- Murphy, T. F. (2012). Double-effect reasoning and the conception of human embryos. *Journal of Medical Ethics*, 39(8), 529-532. doi:10.1136/medethics-2012-100534
- National Academies of Sciences, Engineering, and Medicine. (2017). *Human genome editing: science, ethics, and governance*. National Academies Press.

- Olsthoorn-Heim, E., De Wert, G., Winter, H., te Braake, T., Heineman, M., Middelkamp, A., & Nierse, C. (2006). Evaluatie embryowet. *Den Haag: ZonMw. Reeks evaluatie regelgeving: deel, 20.*
- Parfit, D. (1987). *Reasons and Persons*, reprint with corrections: Oxford: Clarendon Press.
- Pennings, G., de Wert, G., Shenfield, F., Cohen, J., Tarlatzis, B., & Devroey, P. (2007). ESHRE Task Force on Ethics and Law 13: the welfare of the child in medically assisted reproduction. *Human Reproduction*, 22(10), 2585-2588. doi:10.1093/humrep/dem237
- Pennings, G., Segers, S., Debrock, S., Heindryckx, B., Kontozova-Deutsch, V., Punjabi, U., van de Velde, H., van Steirteghem, A., Mertes, H. (2017). Human embryo research in Belgium: an overview. *Fertility and Sterility*, 108(1), 96-107.
- Penzias, A., Bendikson, K., Butts, S., Coutifaris, C., Falcone, T., Fossum, G., Gitlin, S., Gracia, C., Hansen, K., La Barbera, A. (2018). The use of preimplantation genetic testing for aneuploidy (PGT-A): a committee opinion. *Fertility and Sterility*, 109(3), 429-436.
- Read, R., & O'Riordan, T. (2017). The Precautionary Principle Under Fire. *Environment: Science and Policy for Sustainable Development*, 59(5), 4-15.
- Sandin, P. (1999). Dimensions of the precautionary principle. *Human and Ecological Risk Assessment: An International Journal*, 5(5), 889-907.
- Schippers, E. I. (2016). Kabinetsreactie op het Pallas rapport. (829299-140854-PG). Den Haag.
- Shapshay, S., & Pimple, K. D. (2007). Participation in biomedical research is an imperfect moral duty: a response to John Harris. *Journal of Medical Ethics*, 33(7), 414-417.
- Sharpe, R. M. (2018). *Of mice and men: long-term safety of assisted reproduction treatments*: Oxford University Press.
- Sutcliffe, A. G., & Ludwig, M. (2007). Outcome of assisted reproduction. *The Lancet*, 370(9584), 351-359.
- te Braake, T. A. (1998). Wetgeving inzake wetenschappelijk onderzoek met menselijke embryo's. Enkele aandachtspunten. *Tijdschrift voor gezondheidsrecht*, 22(5), 93.
- van Beers, B. (2016, June 2nd 2016). Wijzig de status van het embryo niet buiten het parlement om. *de Volkskrant*, p. 22. Retrieved from <http://www6.lexisnexis.com/publisher/EndUser?Action=UserDisplayFullDocument>

&orgId=102575&Em=1&topicId=158330010&docId=l:2596587146&start=6&md5=183a284e7a79ec69ed4d3f82577d5f0&sendDate=20160602

- Verlinsky Y, Cieslak J, Freidine M, Lvakhnenko V, Wolf G, Kovalinskaya L, White M, Lifchez A, Kaplan B, Moise J. (1995). Pregnancies following pre-conception diagnosis of common aneuploidies by fluorescent in-situ hybridization. *MHR: Basic science of reproductive medicine*, 1(5): 265-269.
- Warren, M. A. (1997). *Moral status: Obligations to persons and other living things*: Clarendon Press.
- Wert, G. d. (1987). In vitro fertilisatie en experimenten met embryo's. Ethisch-filosofische beschouwingen. *Alg. Ned. Tijdschrift v. Wijsbegeerte*, 79, 210-225.
- Winter, H. (2012). *Evaluatie Embryowet en Wet donorgegevens kunstmatige bevruchting*: ZonMw.

Valorization addendum

In this thesis, I investigated how to responsibly deal with medical risks of new reproductive technologies for children born as a result of their clinical application. The results of this study can contribute to improve the innovation and practice of medically assisted reproduction (MAR). In this addendum, I first explain how this research benefits relevant stakeholders. Secondly, I discuss which valorization activities have been undertaken.

Contributing to responsible innovation in medically assisted reproduction

The insights generated from this research are relevant to four groups of stakeholders. First, to MAR families. Responsible innovation in MAR is aimed to improve the quality and safety of the techniques for those dependent on them and for those born after them. This thesis provides insights into how to innovate MAR in an ethically responsible way, thereby contributing to safer reproductive treatment, both in the short and long term.

Second, this research provides handhold for researchers, as it illustrates how different research steps (preclinical research, clinical study, follow-up) can be conducted in practice responsibly. Additionally, it may answer some ethical questions that researchers themselves struggle with, such as questions surrounding the justification of research with animals and/or human embryos.

Third, this thesis provides insights on how healthcare professionals in the field of MAR should deal with offspring safety risks. This counts for both professional organizations, such as the European Society of Human Reproduction and Embryology (ESHRE) or the American Society for Reproductive Medicine (ASRM), and for individual professionals. For example, this thesis proposes to make several amendments to the normative framework of responsible innovation that ESHRE proposes. These amendments are discussed more extensively in the Discussion chapter. As to individual professionals, for example, a justification for motivating patients to participate in follow-up in specific contexts is given.

Last, this thesis hopes to inspire policy makers to adjust legislation and gives practical advice on how to do so. Firstly, legislation that is not ethically sound and hampers responsible innovation in MAR should be changed or abandoned. For example, as argued in the

Discussion, the ban on creating human embryos for research as contained in the relevant legislation of several European countries including the Netherlands, should be lifted on conditions. With this specific recommendation, this thesis contributes to a current national debate about this issue for which the initiative was taken by the Dutch government in the fall of 2019. And secondly, this thesis insists on the need for a legal framework to make sure self-regulation is followed and treatments are efficient and safe.

Undertaken valorization activities

I have undertaken multiple valorization activities to generate attention for the findings of this study. First, the results of my work have been published in international journals focusing on technical and normative aspects of reproductive and genetic medicine. This way, the outcomes of this research were available for both clinicians and the scientific community.

Second, I published two brief Dutch-language articles. One considered the Dutch Embryos Act and was published in a Dutch newspaper¹. In this article, I argued that there is no moral justification for upholding the current categorical ban on creating human embryos for research and that this ban should be lifted, because it hampers responsible innovation. The aim was to bring this topic to the attention of the general public and policy makers. The second article was published in the journal of the Dutch Society for Obstetrics & Gynaecology (NTOG).² This gave me the opportunity to address a broad range of possibly interested professionals with the main findings of my paper on innovation and precaution (Chapter 4 of this thesis). The NTOG-article was written on invitation for a special section of the journal devoted to the presentations by Dutch professionals given at the 2019 ESHRE conference in Vienna.

Third, I presented my work at three annual ESHRE conferences (2017, 2018 and 2019). This way, the results of my research were made available to a large, international, group of clinicians, scientists and ethicists in the field of reproductive medicine. I was invited

¹ Verna Jans en Wybo Dondorp. Sta weken voor onderzoek toe. *Trouw* 14 juni 2017; <https://www.trouw.nl/nieuws/sta-kweken-embryo-voor-onderzoek-toe~ba2f7d74/>

² V Jans, WJ Dondorp, GMWR de Wert. Verantwoorde introductie van nieuwe voortplantingstechnieken. Tussen innovatie en voorzorg. *Ned Tijdschr Obstet Gynaecol* 2020; 133 (4): 205-206. http://www.ntog.nl/dynamic/media/3/documents/NTOG_2020_4_web.pdf

as a speaker at a pregress-course on the safety of MAR preceding ESHRE's 2020 conference in Copenhagen. This would have been a further occasion to bring my work under the attention of the most relevant group of professionals. Unfortunately, due to the corona-pandemic, this conference was cancelled.

Last, this thesis is part of the Science and Ethics of stem cell derived Gametes (SEGa) project. I presented during public events organized by the project's team. For example, in 2020, a science meeting discussing the outcomes of the project was organized in Ghent, for which people could register. This way, my research outcomes were also brought to the attention of societal partners of the project, such as representatives of patient organizations, and to a lay audience.

Samenvatting

Sinds de geboorte van de eerste reageerbuisbaby, ongeveer veertig jaar geleden, is de ontwikkeling van voortplantingstechnologie snel gegaan. Een wervelwind van innovaties heeft mensen die afhankelijk zijn van medisch geassisteerde voortplanting veel opties gebracht. Het werkveld van voortplantingsgeneeskunde heeft echter ook kritiek gekregen. Dit betreft met name de manier waarop nieuwe technieken worden geïntroduceerd in de kliniek. Volgens critici zou dit op een onzorgvuldige en onverantwoorde manier gebeuren, waardoor er onnodig medische risico's ontstaan voor patiënten en kinderen. Tegen deze achtergrond stel ik de vraag hoe er verantwoord omgegaan kan worden met medische risico's voor kinderen, die het introduceren van nieuwe reproductieve technieken met zich meebrengt. Ik onderzoek deze vraag in dit proefschrift door gebruik te maken van zowel theoretische kaders als empirische bevindingen.

In *hoofdstuk een* leid ik het onderwerp van mijn studie in en bespreek ik de methodologie. De introductie van nieuwe voortplantingstechnieken gaat vaak op basis van *trial and error*: al doende moet blijken wat wel en wat niet werkt. Dat is heel anders dan bij bijvoorbeeld de introductie van nieuwe medicijnen, waarbij vooraf grondig onderzoek plaatsvindt. Bij de introductie van nieuwe voortplantingstechnieken zijn de veiligheidsmaatregelen, zoals preklinisch effectiviteits- en veiligheidsonderzoek, vaak onvoldoende. Hoewel follow-up onderzoek niet wijst op ernstige negatieve gezondheidseffecten als gevolg van het gebruik van reproductieve technieken, wil dat niet zeggen dat de manier waarop nieuwe voortplantingstechnieken nu naar de kliniek worden gebracht geen verbetering zou behoeven. Er zijn potentieel riskante technieken in ontwikkeling, zoals de toepassing van kiembaanmodificatie voor voortplantingsdoeleinden of het gebruik van gameten gecreëerd uit stamcellen. Dat maakt het des te belangrijker om mogelijke gezondheidsrisico's uiterst serieus te nemen.

De European Society of Human Reproduction and Embryology (ESHRE) heeft een normatief framework voorgesteld voor verantwoorde innovatie op het gebied van medisch geassisteerde voortplanting. Dit framework bestaat uit drie opeenvolgende onderzoeksfasen: (1) preklinisch onderzoek, gebruik makende van dieren en menselijke

embryo's; (2) het introduceren van nieuwe technieken in een klinische onderzoekssetting met registratie van uitkomsten en (3) langdurig follow-up onderzoek. Het is momenteel nog onduidelijk hoe dit normatieve framework in de praktijk te gebruiken en hoe de drie onderzoeksfasen op een ethisch verantwoorde wijze uit te voeren. Dit proefschrift geeft hier invulling aan. Het bevat een ethische toetsing van de huidige innovatiepraktijk van medisch geassisteerde voortplanting en geeft bovendien aanbevelingen voor een verantwoorde introductie van nieuwe voortplantingstechnieken.

Om de centrale vraag te beantwoorden heb ik zowel literatuuronderzoek als empirisch onderzoek uitgevoerd. De literatuur bevatte wetenschappelijke en ethische papers over voortplantings- en genetische geneeskunde, naast boeken en internationale overheidsdocumenten over debat en wetgeving rondom voortplantingsgeneeskunde. Om meer inzicht te krijgen in de introductie van de vele voortplantingstechnieken die aan de orde komen in dit proefschrift, zijn er meerder individuele interviews gehouden met experts op het gebied van medisch geassisteerde voortplanting. Verder hebben er twee focusgroepbijeenkomsten plaatsgevonden; de ene met wetenschappers, klinici en ethici en de andere met individuen of koppels die ervaring hadden met medisch geassisteerde voortplanting. Voor de ethische analyse in dit proefschrift is gebruik gemaakt van *Wide Reflective Equilibrium*, een methode die zowel normatieve en achtergrondtheorieën als bevindingen en intuïties uit de praktijk meeneemt.

Preklinisch onderzoek, waaronder dieronderzoek, wordt steeds vaker gezien als cruciaal onderdeel van de verantwoorde introductie van nieuwe voortplantingstechnieken. Tegelijkertijd klinkt in de maatschappij en de politiek een steeds luidere oproep om minder dieronderzoek uit te voeren. Deze twee ontwikkelingen lijken haaks op elkaar te staan. Daarom onderzoek ik in *hoofdstuk twee* of het gebruik van dieren in preklinisch onderzoek van nieuwe voortplantingstechnieken moreel acceptabel is en zo ja, onder welke voorwaarden. Vertrekkende vanuit de premisse dat de morele status van dieren niet vergelijkbaar is met die van mensen, beargumenteer ik dat het instrumentele gebruik van dieren voor onderzoek moreel acceptabel is onder de voorwaarden van subsidiariteit en proportionaliteit. Subsidiariteit houdt in dat de voordelen van dieronderzoek niet via andere, moreel minder problematische wegen te verkrijgen zijn. Ik leg uit dat het onaannemelijk is dat alternatieve modellen het dieronderzoek naar nieuwe voortplantingstechnieken geheel

kunnen vervangen. Modellen zoals het gebruik van micro-organismen en *in vitro* cel- en weefselculturen van dieren of mensen kunnen namelijk geen informatie geven over de ontwikkeling van embryo's en foetussen, noch over de gezondheid van (multigenerationele) nakomelingen. Preklinisch dieronderzoek naar nieuwe voortplantingstechnieken voldoet dus aan de voorwaarde van subsidiariteit.

Proportionaliteit houdt in dat de mate van inbreuk op het welzijn van dieren in verhouding moet zijn met de voordelen van dieronderzoek. De belasting voor dieren in onderzoek naar nieuwe voortplantingstechnieken is laag vergeleken met andere geaccepteerde vormen van dieronderzoek, zoals in kanker- of brandwondenstudies, of met het gebruik van dieren in de vleesindustrie. De voordelen van dieronderzoek naar nieuwe voortplantingstechnieken beoordeel ik op twee niveaus. Ten eerste: is de waarde van medisch geassisteerde voortplanting belangrijk genoeg om het gebruik van dieren voor onderzoek te rechtvaardigen? Het voornaamste doel is hier om mensen te helpen een genetisch eigen kind te krijgen. Ook al lopen opvattingen over het belang van de genetische band sterk uiteen, het lijkt niet juist om te zeggen dat dit slechts een triviale voorkeur is. Aangezien er verschillende opvattingen zijn over wat een betekenisvol leven inhoudt en teneinde deze diversiteit van opvattingen te respecteren, kan beargumenteerd worden dat de praktijk van medisch geassisteerde voortplanting inderdaad een moreel waardevol doel dient. Als we dit vergelijken met de algemeen aanvaarde opvatting dat het gebruik van dieren voor voedsel acceptabel kan zijn onder (in de praktijk van de bio-industrie maar al te vaak geschonden) voorwaarden, lijkt onderzoek gericht op het verbeteren van medisch geassisteerde voortplanting door het preklinisch onderzoeken van nieuwe voortplantingstechnieken voldoende belangrijk om ook voorwaardelijk gebruik van dieren voor dat doel te accepteren. Ten tweede: is het doel van *onderzoek* naar nieuwe voortplantingstechnieken belangrijk genoeg om dierstudies te rechtvaardigen? Ik stel dat professionals die voortplantingstechnieken aanbieden een gedeelde verantwoordelijkheid hebben voor het welzijn van hun patiënten en de kinderen die zij helpen tot stand te brengen. Onder die verantwoordelijkheid valt ook het testen van de effectiviteit en veiligheid van nieuwe technieken om onnodige medische risico's te beperken. Het belang van dit doel weegt zwaarder dan de beperkte inbreuk op het welzijn van dieren in onderzoek naar nieuwe voortplantingstechnieken.

Het gebruik van dieren voor preklinisch onderzoek naar nieuwe voortplantingstechnieken voldoet aan de voorwaarden van subsidiariteit en proportionaliteit en is dus, in principe, gerechtvaardigd. In principe, omdat de rechtvaardiging altijd beoordeeld moet worden op per specifiek geval, aangezien de proportionaliteit van dieronderzoek in de praktijk nog steeds afhankelijk is van de specifieke kenmerken van dat onderzoek. Verder blijft het belangrijk, aangezien dieronderzoek onvermijdelijk een morele prijs met zich meebrengt, om de noodzaak van dierproeven zoveel mogelijk te beperken en te streven naar het ideaal van de 3V's (*vervanging* van dierproeven door andere methodes; *vermindering* van het aantal proefdieren zonder in te boeten op informatie; *verfijning* van methodes zodat het welzijn van de dieren verbetert). Hoewel volledige vervanging van dieronderzoek naar nieuwe voortplantingstechnieken niet realistisch is, is het desalniettemin van groot belang om na te denken over mogelijke alternatieven, waaronder het gebruik van menselijke embryo's voor onderzoeksdoelinden.

Onderzoek met menselijke embryo's als alternatief voor dierproeven staat centraal in *hoofdstuk drie*. Is er een ethisch te prefereren volgorde van preklinische onderzoeksstappen in de ontwikkeling en introductie van voortplantingstechnieken? De ESHRE-aanbevelingen stellen dat preklinisch veiligheidsonderzoek, inclusief het gebruik van diermodellen en menselijke embryo's, een vereiste is voor verantwoorde innovatie. De morele gevoeligheid van onderzoek op menselijke embryo's heeft echter geleid tot voorschriften en richtlijnen die studies met menselijke embryo's verbieden wanneer onderzoek niet ook met dieren kan worden uitgevoerd. Ik noem dit de gebruik-eerst-dieren-regel, in het Engels '*Use Animals First*' ofwel de UAF-regel. In onderzoek naar nieuwe voortplantingstechnieken vertaalt zich dit in de notie van een 'ideale keten' van preklinische onderzoeksstappen. Onderzoek met menselijke embryo's is alleen dan te overwegen wanneer eerst veelbelovende resultaten eerst bij dieren zijn verkregen, of wanneer het duidelijk is dat dierstudies niets kunnen bijdragen aan het beantwoorden van de onderzoeksvraag.

Vanuit zuiver wetenschappelijk perspectief bestaat er geen algemeen geprefereerde volgorde van onderzoeksstappen. De optimale manier om een onderzoeksvraag te beantwoorden ligt immers niet van tevoren vast, maar wordt telkens opnieuw bepaald. De uiteindelijke volgorde kan met de UAF-regel overeenstemmen of ervan afwijken. In de praktijk blijkt die regel echter vaak strikter geïnterpreteerd te worden, wat resulteert in de

‘ideale keten’. Het wordt moreel problematisch wanneer ethische onderzoekscommissies vervolgens eisen dat aan menselijk-embryo-onderzoek altijd dieronderzoek vooraf dient te gaan, ook wanneer het dieronderzoek naar verwachting van gering nut is. Verspilling van onderzoeksgeld en -tijd kan het gevolg zijn van deze aanpak, die er ook toe kan leiden dat de voordelen van dieronderzoek niet meer opwegen tegen de inbreuk van het welzijn van de gebruikte dieren.

Ondanks zijn potentiële wetenschappelijke en ethische nadelen blijkt de UAF-regel vaak van toepassing te zijn. Ik ga na welke morele argumenten hieraan mogelijk ten grondslag liggen en weerleg ze. Een eerste mogelijk moreel argument om de UAF regel te behouden is dat menselijke embryo’s in wetenschappelijk onderzoek ‘meer te verliezen hebben’ dan dieren en dat zij daarom als onderzoeksmateriaal de minst gewenste optie zijn. Ik weerleg dit argument als volgt: menselijke preïmplantatie-embryo’s kunnen in tegenstelling tot dieren geen pijn of psychologisch leed ervaren en dus heeft onderzoek geen effect op hun welzijn. Ook een voortijdige dood als vorm van inbreuk op het welzijn vormt geen argument ten faveure van de UAF-regel. Voor zover dieren een bepaalde psychologische connectie hebben tot zichzelf en tot hun toekomstige leven, kan het doden van dieren een inbreuk op hun welzijn vormen. In het geval van menselijke embryo’s is deze psychologische connectie afwezig.

Een tweede mogelijk moreel argument is dat menselijke embryo’s, in tegenstelling tot dieren, een speciale morele status zouden hebben, ongeacht of zij belangen omtrent hun welzijn hebben of niet. Ik beargumenteer dat dieren in het algemeen onder de grens vallen van wat het inhoudt een persoon te zijn, wat betekent dat instrumenteel gebruik onder voorwaarden acceptabel is. Ik stel dat menselijke embryo’s meestal een beperkte morele status toegeschreven krijgen, die toeneemt naarmate het embryo zich verder ontwikkelt. Deze status is relatief laag in vergelijking tot personen, maar niet tot dieren, aan wie zo’n morele status helemaal niet kan worden toegekend. Degenen die de UAF-regel verdedigen zouden dat op grond van dit verschil kunnen doen.

Het derde en laatste morele argument betreft de belangen van eiceldonoren in de gevallen waar het niet mogelijk is embryo’s te gebruiken die zijn overgebleven na een IVF-behandeling, de zogeheten restembryo’s. Ik leg uit dat eiceldonatie voor wetenschappelijk onderzoek mogelijke risico’s kan opleveren voor donoren, zonder dat zij er direct voordeel

van hebben. Aangezien het echter mogelijk is om fysieke risico's te vermijden en sociale risico's te minimaliseren, beargumenteer ik dat deze uitdagingen niet dusdanig groot zijn dat ze doorslaggevend zijn om de UAF-regel te behouden.

Ik concludeer dat de UAF-regel gebaseerd lijkt op de opvatting dat menselijke embryo's een bijzondere morele status hebben, zonder het welzijnspectief in acht te nemen. Wanneer dat wel wordt meegenomen, is het minder duidelijk of de UAF-regel moreel houdbaar is. De kernboodschap blijft dat de stap naar de kliniek pas verantwoord is na adequate preklinische veiligheidsstudies. De UAF-regel kan die kernboodschap in de weg staan, echter zonder dat daar gegronde morele argumenten voor zijn. Mijn advies is om de UAF-regel in wettelijke documenten en zelfregulering te vervangen door een normatief framework dat dierstudies niet *a priori* verkiest boven het gebruik van menselijke embryo's voor onderzoek. Dit houdt ook in dat de notie van een 'ideale keten' in onderzoek naar nieuwe voortplantingstechnieken aanpassing behoeft. In lijn hiermee geldt bovendien dat ethische onderzoekscommissies, in protocollen voor het gebruik van menselijke embryo's in preklinisch onderzoek, geen voorafgaande dierstudies moeten eisen die naar verwachting niet (voldoende) informatief zijn of die hogere morele kosten met zich meebrengen dan embryo-onderzoek.

Zoals hierboven al kort genoemd, is het werkveld van medisch geassisteerde voortplanting vaak bekritiseerd voor het introduceren van nieuwe voortplantingstechnieken zonder systematisch preklinisch onderzoek naar de werkzaamheid en veiligheid en zonder follow-up van kinderen die met die technologieën zijn verwekt. In *hoofdstuk vier* beoordeel ik of deze kritiek gerechtvaardigd is en zo ja, hoe het werkveld bij nieuwe technieken dan wel moet omgaan met de veiligheidsrisico's voor kinderen bij het introduceren van nieuwe technieken. Voor de vraag of de kritiek gerechtvaardigd is, heb ik wat betreft de nakomelingen onderzocht of veiligheidsoverwegingen van nakomelingen al dan niet een rol speelden in de geschiedenis van de introductie van nieuwe voortplantingstechnieken. Ik heb hiervoor drie *casestudy's* verricht. Twee *casestudy's* zijn uit het verleden: intracytoplasmatische sperma-injectie (ICSI) en pre-implantatie genetische diagnostiek (PGT). De derde is hedendaags: mitochondriale vervangingstherapie (MRT). Als kader voor de beoordeling gebruikte ik de vier dimensies van het omgaan met onzekere risico's, door Per Sandin geformuleerd in zijn analyse van de structuur van het zogeheten

voorzorgsbeginsel. Als er een bepaalde (1) dreiging is, die (2) onzeker is, dan is er een bepaalde (3) actie (4) vereist.

Uit de casestudy's blijkt dat er inderdaad aandacht is besteed aan de eerste twee dimensies. Zorgen over de veiligheid van nieuwe technieken stonden op de debatagenda en leidden soms tot voorstellen voor concrete veiligheidsmaatregelen (bijvoorbeeld parameters voor spermaselectie bij ICSI). Tegelijkertijd tonen de casestudy's ook aan dat de stap van dimensie twee naar drie (en vier) niet vanzelfsprekend is. Er kan een spectrum geïdentificeerd worden van hoe om te gaan met veiligheidsrisico's voor nakomelingen. Aan de ene kant van het spectrum ligt de 'sterke' interpretatie van het voorzorgsbeginsel: voer bepaalde technieken niet in zolang er onzekerheid bestaat over de veiligheidsrisico's voor nakomelingen. Deze redenering lijkt bijvoorbeeld ten grondslag te liggen aan het Nederlandse moratorium op MESA en TESE, twee chirurgische varianten van ICSI. Het verbod op deze varianten van ICSI werd pas in 2012 (MESA) en 2014 (TESE) opgeheven, nadat er geruststellende veiligheidsgegevens beschikbaar waren gekomen. Aan de andere kant van het spectrum ligt het innovatieprincipe. In tegenstelling tot het voorzorgsbeginsel, waarbij onzekerheid tot actie moet leiden, benadrukt het innovatiebeginsel dat alleen bekende schadelijkheid een reden kan zijn om de toepassing van nieuwe technologieën te beperken. Voorbeelden hiervan zijn het huidige gebruik van ICSI voor niet-mannelijke onvruchtbaarheidsproblemen, PGT-A om de kans op zwangerschap te vergroten, en MRT in de context van 'eicelverjonging', zonder bewezen effectiviteit. Tussen de uitersten in liggen kunnen 'lichtere' vormen van het voorzorgsprincipe geïdentificeerd worden. Hier wordt onzekerheid nog steeds beschouwd als een geldige reden om maatregelen te nemen, maar deze maatregelen hoeven vervolgens niet restrictief te zijn.

Ik concludeer dat aanbevelingen uit de beroepsgroep doorgaans worden opgevat als vrijblijvende suggesties en niet als richtinggevend voor gedeelde professionele verantwoordelijkheid. Vanuit commercieel oogpunt is dit misschien geen probleem. Echter, voor een groot deel van het reproductieve veld, lijkt de benadering in lijn met het innovatieprincipe in strijd met het idee dat professionals in geassisteerde voortplanting de verantwoordelijkheid hebben om rekening te houden met het welzijn van de kinderen die zij helpen te maken. Omgekeerd lijken 'sterke' vormen van het voorzorgsbeginsel lastig te rechtvaardigen vanwege de moreel zwaarwegende belangen van degenen die afhankelijk zijn

van medische hulp bij het krijgen van kinderen. Ik stel daarom een middenpositie voor, die zich kenmerkt door een beleid van systematische voorzorg. Dit betekent dat veiligheidsrisico's verbonden aan de introductie van nieuwe voortplantingstechnieken zoveel mogelijk moeten worden geminimaliseerd. De aanbevelingen die ESHRE heeft gedaan voor verantwoord innoveren in medisch geassisteerde voortplanting (i.e. preklinisch onderzoek, klinisch onderzoek, follow-up onderzoek) kunnen hiervoor als voorbeeld dienen. Verder zouden overheden een belangrijkere rol moeten innemen door juridische kaders te formuleren voor een betere naleving van de aanbevelingen in de beroepspraktijk.

Volgens het framework van ESHRE is langdurig follow-up onderzoek essentieel voor het verantwoord introduceren van voortplantingstechnieken. Hoe kan dit follow-uponderzoek op een moreel verantwoorde wijze uitgevoerd worden? In *hoofdstuk vijf* heb ik twee van de belangrijkste ethische overwegingen hieromtrent onderzocht: de proportionaliteit van de onvermijdelijke lasten en risico's voor deelnemende gezinnen, en de implicaties van het beginsel van respect voor autonomie.

Follow-up onderzoek bij medisch geassisteerde voortplanting betreft het monitoren van de gezondheid van zowel patiënten als nakomelingen in de periode na behandeling. In dit proefschrift heb ik mij uitsluitend gericht op follow-up onderzoek bij nakomelingen, omdat dat uniek is aan de follow-up van voortplantingstechnologie. Ik leg uit dat follow-up (zowel kort- als langdurend) van kinderen die geboren zijn na toepassing van medisch geassisteerde voortplanting belangrijk is voor vier verschillende groepen individuele stakeholders: 1) voor die kinderen zelf, 2) voor hun nakomelingen, 3) voor toekomstige kinderen geboren vanuit die techniek en 4) voor ouders en de samenleving als geheel. Enkele IVF-centra hebben een substantiële bijdrage geleverd aan follow-up onderzoek van voortplantingstechnieken, maar zij zijn nog steeds de uitzondering.

Om de proportionaliteitsvraag te beantwoorden onderzoek ik of de mogelijke voordelen van follow-up opwegen tegen de mogelijke risico's en lasten voor de deelnemers. Over het algemeen bestaat follow-up uit non-invasieve tests, waardoor de medische risico's beperkt zijn. Verder is het mogelijk dat follow-up leidt tot psychologische belasting. Omdat follow-up onderzoek van voortplantingstechnieken voornamelijk onderzoek met minderjarigen betreft, is het van belang rekening te houden met speciale voorwaarden die voor dergelijk onderzoek gelden. Zo wordt onderzoek met minderjarigen als acceptabel

beschouwd wanneer er potentieel voordeel is voor henzelf, de risico's worden geminimaliseerd en het voordeel de risico's compenseert. Is er geen duidelijk voordeel, dan is onderzoek met minderjarigen alleen toegestaan als 1) antwoord op de onderzoeksvraag alleen mogelijk is door onderzoek te doen met minderjarigen (criterium van groepsgebondenheid); en 2) de lasten en risico's voor die minderjarigen minimaal zijn. Omdat alleen onderzoek met kinderen inzicht kan geven in de mogelijke impact op de ontwikkeling van kinderen die zijn verwekt met voortplantingstechnologie, voldoet follow-up van voortplantingstechnieken aan het criterium van groepsgebondenheid. Toch hangt het niveau van risico's en lasten, en daarmee de aanvaardbaarheid van follow-up bij minderjarigen, af van de specifieke onderzoeksvraag en de specifieke onderzoeksfase. Ik stel vast dat dit leidt tot een lastig dilemma: hoewel het belang van follow-up studies het grootst is in de experimentele fase van nieuwe voortplantingstechnieken, zijn de risico's en lasten hier het hoogst.

Wanneer een specifieke follow-upstudie proportioneel bevonden is, is het vervolgens belangrijk om rekening te houden met het beginsel van respect voor autonomie. Dit betekent dat deelname aan onderzoek altijd vrijwillig moet zijn en op basis van goede informatie. Bij follow-up onderzoek met minderjarigen zullen hun ouders deze beslissing moeten nemen. Echter, als deelname aan onderzoek altijd vrijwillig moet zijn, wat betekent dat dan voor de aanvaardbaarheid van het actief motiveren van ouders om hun kinderen te laten deelnemen? Het hangt er vanaf of er sprake is van een bepaalde morele verantwoordelijkheid bij ouders en kinderen om deel te nemen aan het onderzoek. In dat geval zou te verdedigen zijn dat professionals hen hierop mogen aanspreken. Ik noem twee mogelijke argumenten voor het toeschrijven van een dergelijke verantwoordelijkheid aan ouders van kinderen verwerkt via medisch geassisteerde voortplanting. Ten eerste, in het geval van experimentele of innovatieve behandelingen hebben ouders een conditionele verantwoordelijkheid ten opzichte van hun kinderen om hen mee te laten doen in follow-up onderzoek dat gezondheidsproblemen vroegtijdig kan opsporen. Ten tweede, deelname aan follow-up onderzoek kan gezien worden als een manier om iets terug te doen voor de voordelen die medisch geassisteerde voortplanting hun en hun kinderen (die er hun bestaan aan danken) gebracht heeft. Het morele argument hier is 'reciprociteit'. Ik ga niet zover als John Harris, volgens wie er een op reciprociteit gebaseerde algemene (voor iedereen geldende) plicht

bestaat om bij te dragen aan medisch-wetenschappelijk onderzoek ter verbetering van de gezondheidszorg. Hij beschouwt gezondheidszorg als een sociale praktijk waar ieder lid van een moderne westerse samenleving zelf ook de voordelen van ondervindt. Terecht is opgemerkt dat er hoogstens zoiets bestaat als een voor ieder geldende indirecte morele plicht om bij te dragen aan 'the general good'. Bijdragen aan medisch- wetenschappelijk onderzoek (als proefpersoon of donor) is slechts een van vele manieren om zich te kwijten van de verantwoordelijkheid om 'iets terug te doen', naast het doen van vrijwilligerswerk etc. Uitgaande van zo'n indirecte plicht, maakt het echter wel uit dat voor deze specifieke bijdrage (follow-up) alleen diegenen in aanmerking gevraagd kunnen worden die de onderzochte procedure zelf hebben ondergaan. Mijn stelling is dat van hen verwacht mag worden dat zij een verzoek tot deelname aan follow-up om die reden ten minste serieus zullen overwegen.

Gezien het belang van follow-uponderzoek voor verantwoorde innovatie in medisch geassisteerde voortplanting moet het werkveld een actievere houding aannemen om dit onderzoek uit te voeren. Het is professionals moreel geoorloofd om families actief te motiveren om deel te nemen, op voorwaarde dat de studie voldoet aan het proportionaliteitsbeginsel. In het geval van formeel klinisch onderzoek naar nieuwe voortplantingstechnieken mogen professionals zelfs een stap verder gaan. Zij mogen deelname aan follow-uponderzoek dan als (niet-afdwingbare) voorwaarde stellen voor toegang tot de behandeling.

In *hoofdstuk zes* bespreek ik de bredere implicaties van de conclusies van de voorgaande hoofdstukken. Ik verdeel dit in vier subthema's: dieronderzoek, onderzoek met menselijke embryo's, de praktische implicaties voor het normatieve framework voor verantwoord innoveren in medisch geassisteerde reproductie en tot slot de standaard voor de beoordeling van veiligheid. Bij het eerste subthema onderzoek ik of er moreel een verschil te maken is tussen de gebruikte diersoorten in preklinisch onderzoek. Ik onderscheid twee mogelijke redeneringen. Enerzijds kan er een verschil gemaakt worden gebaseerd op de verschillende complexiteit van het mentale en sociale leven van de betreffende diersoorten, zonder daarbij een verschil in morele status toe te schrijven aan die dieren. De inbreuk op het welzijn van dieren in dieronderzoek kan per diersoort verschillen, omdat elke soort zijn eigen behoefte heeft. Zo vormt het hogere niveau van psychologische en sociale complexiteit bij primaten een grotere morele barrière om hen te gebruiken voor onderzoek dan bij andere

diersoorten het geval is. Vanuit die redenering is onderzoek met, bijvoorbeeld, varkens dus moreel uitdagender dan onderzoek met muizen. Anderzijds zou een verschil gemaakt kunnen worden op basis van de opvatting dat alle diersoorten, juist vanwege de variaties in complexiteit, een verschil in morele status hebben. Dit zou leiden tot een andere vorm van redeneren, waarbij bijvoorbeeld onderzoek met varkens problematischer is dan met muizen, omdat varkens een hogere morele status zouden hebben. Ongeacht welke van de mogelijke redeneringen gevolgd wordt, is het niet vanzelfsprekend dat de voorkeur moet uitgaan naar dieren die simpelere soortspecifieke behoeften of een lagere morele status hebben. Voor bepaalde onderzoeksvragen kan het gebruik van dieren met uitdagendere soortspecifieke behoeften een duidelijke toegevoegde waarde hebben, die niet bereikt kan worden met andere diersoorten. In het licht van de verantwoordelijkheid die professionals in medisch geassisteerde voortplanting dragen voor hun patiënten en hun kinderen beargumenteer ik het volgende: onderzoek met dieren die uitdagendere soortspecifieke behoeften hebben kan, in specifieke gevallen, meer proportioneel zijn dan onderzoek met andere diersoorten. Het is derhalve van belang om kritisch te kijken naar de vooroordelen betreffende de proportionaliteit van het gebruik van complexere versus minder complexe diersoorten voor onderzoek en om ruimte te laten voor verandering. Dit houdt in dat er, onder voorwaarden, ruimte moet zijn om meer complexe diersoorten te gebruiken voor preklinisch onderzoek in zoverre dit gebruik duidelijk voordeel biedt.

Bij het tweede subthema onderzoek ik of er een moreel verschil gemaakt kan worden tussen het gebruiken van menselijke restembryo's en het speciaal tot stand brengen van embryo's voor preklinisch onderzoek (onderzoeksembryo's). Restembryo's zijn al enkele dagen oud, wat hun toepassing voor onderzoek in medisch geassisteerde voortplanting beperkt. Voor onderzoek naar de effectiviteit en veiligheid van bepaalde nieuwe voortplantingstechnieken, zoals het invriezen van eicellen, is het bovendien nodig om embryo's te onderzoeken die gecreëerd zijn met de onderzochte technologie. Het gebruik van speciaal tot stand gebrachte embryo's kan dus een grote bijdrage leveren aan het testen van nieuwe voortplantingstechnieken. Ik laat zien dat moreel de argumenten voor een moreel onderscheid tussen het gebruik van restembryo's en onderzoeksembryo's veelal berust op het argument dat de twee soorten embryo's met verschillende intenties gecreëerd zijn. Het tot stand brengen van onderzoeksembryo's zou het menselijk leven reduceren tot louter

gebruiksmiddel, terwijl embryo's die overblijven na IVF gecreëerd zouden zijn met de intentie uit te groeien tot kind. Vervolgens beargumenteer ik dat, in een maatschappij die reguliere IVF-behandelingen als moreel acceptabel beschouwt, het intentie-argument niet steekhoudend is om bij onderzoekstoepassingen een moreel verschil te verdedigen tussen restembryo's en onderzoeksembryo's. Reguliere IVF-behandelingen reduceren embryo's immers evenzeer tot gebruiksmiddel, aangezien het aantal gecreëerde embryo's bewust hoger is dan het aantal teruggeplaatste embryo's in de baarmoeder, met het doel zwangerschapskansen te vergroten. Verder bespreek ik de risico's voor eiceldonoren. Deze risico's verdienen serieuze aandacht. Ook zijn ze een (additioneel) argument voor het gebruik van restembryo's voor onderzoek wanneer dit dezelfde informatie kan geven als met gebruik van restembryo's. Echter, deze risico's geven niet de doorslag om een categorisch verbod te verdedigen op het speciaal tot stand brengen van embryo's voor onderzoek. Ik concludeer daarom dat het categorische verbod op het maken van embryo's voor onderzoeksdoeleinden opgeheven moet worden, gezien het belang van studies met dergelijke embryo's voor de verantwoorde introductie van voortplantingstechnieken,

Als derde laat ik zien dat dit proefschrift redenen geeft tot drie aanpassingen van het huidige normatieve framework voor verantwoord innoveren in het domein van medisch geassisteerde voortplanting. Ten eerste is meer studie nodig naar hoe om te gaan met enerzijds de noodzaak van transgenerationeel follow-uponderzoek en anderzijds de praktische en ethische uitdagingen die dit onderzoek met zich meebrengt (zoals gebrek aan financiële middelen, privacy, etc.). Ten tweede, de aanbeveling dat gebruik van menselijke embryo's voor preklinisch onderzoek pas overwogen mag worden wanneer dieren ongeschikt zijn, moet worden afgeschaft. Het is essentieel om per geval te bekijken welk onderzoekstraject het meest geschikt is. Ten derde is (inter)nationale wetgeving nodig om te bevorderen dat aanbevelingen vanuit de beroepsgroep opgevolgd worden en vruchtbaarheidsbehandelingen effectief en veilig zijn. De beroepsgroep heeft geen uitvoerende macht, waardoor aanbevelingen in de praktijk te weinig navolging krijgen. Een systeem zoals in het Verenigd Koninkrijk, waar de Human Fertilisation and Embryology Authority (HFEA) ervoor zorgt dat vruchtbaarheidsklinieken en onderzoekscentra aan de wet voldoen, kan een voorbeeld zijn van hoe die wetten te handhaven in IVF-klinieken.

Als laatste stel ik de vraag wat we in medisch geassisteerde voortplanting als een 'acceptabel risico' kunnen beschouwen. Is de evaluatiestandaard die ESHRE momenteel hanteert (hoog risico op ernstige schade) wel voldoende toegespitst op de context van verantwoorde innovatie in voortplantingstechnieken? Tegelijkertijd is een 'zero-risk'-evaluatiestandaard onmogelijk, aangezien alle voortplanting risico's met zich meebrengt. Meer debat over de vraag 'hoe veilig is veilig genoeg?' is noodzakelijk. Verschillende stakeholders, waaronder de beroepsgroep en patiëntorganisaties, moeten hier een rol in spelen.

Acknowledgements

There is one person who started these whole shenanigans: my high school teacher, Mr. Jacobs. Thank you for never believing in me. For your self-confirmed 'heart of stone'. For trying to convince me to drop out of philosophy class, because for that reason, I realized I actually wanted to stay.

In the process of writing this thesis, I got the opportunity to meet many inspiring people. People whom I would like to thank for the various ways in which each of them contributed to my project.

First and foremost, I want to thank my promoters. Wybo Dondorp, for keeping my feet on the ground and continuously challenging me to write precisely and concisely. For your honest and practical feedback. And for your (sometimes questionable) navigation skills during many events and conferences. Yet, one way or another, we always ended up safe at the hotel. Guido de Wert, for keeping the bigger picture in mind, and for always bringing up new inspiring ideas and perspectives to the table. For your constructive feedback on my papers. And for being an interesting conversational partner during the long train rides to SEGa meetings. Keep on your dancing shoes during the ESHRE parties! Heidi Mertes, my co-promotor, for being my faithful co-author and for your challenging comments on the many drafts I sent you. You sure made many SEGa events and conferences a lot more fun.

I would also like to thank all my colleagues from the SEGa project. Not only for sharing your scientific insights, which really sparked my own research, but also for the continuously good times we had at meetings and events. A special thanks to Guido Pennings, Seppe Segers, Ellen Goossens, Ela Kacin, Swati Mishra and Mieke Geens, for being co-author, for being good company, or for being both.

To everyone I got the chance of interviewing or co-authoring with: thank you for sharing your knowledge and experiences with me. Thanks to you, this thesis is built upon both theory and practice. I am especially grateful to all the scientists and clinicians that took their time and effort to show me around their labs and clinics. As an ethicist, I was used to spending a lot of time doing literature research but, thanks to you, I also got to experience

what I did (and did not) read about firsthand. These experiences really enriched my perspective on the practice of medically assisted reproduction.

Thank you to all my (past) colleagues at HES. You brightened up the department with lunches, drinks and outings. In particular, I thank my amazing paranymphs: the two An(n)as. Anna Wolters, being the eloquent and classy lady you are, it truly is a delight to be in your presence. Thank you for our many visits to Lumière and our discussions on movies and life. Ana Pereira Daoud, asking you to organize the department's outing with me was one of the best decisions I ever made. Thank you for all the 'tuna piadina'-lunches, talks about our cats and Netflix series and for always believing in me. Anna and Ana, I could not wish for four better shoulders to support me during my PhD trajectory and defense and I consider you both very good friends. Another special thanks to Greg Stapleton, my beloved office mate. For your interesting perspectives in ethics and the world around us. You sure talked my head off, but I was happy to listen. Thanks for being my wingman during this whole adventure. I also want to explicitly thank Lucie Bastiaens, Sanne van der Hout, Rob Houtepen, Ricky Janssen and Sanne Raap, for your various life-saving moments during these four years. Moreover, thank you Angelique Heijnen, Helen Heutz and Kyra Kramer, for your excellent support and guidance in all things bureaucratic.

I would also like to thank several people for their unconditional mental support, which I think is key for surviving a PhD in general. I would not go as far as to tell you that, without them, I would now be lying in a ditch somewhere; but rocking a fetus position in a corner of my workroom would be a fair possibility.

I thank the 'swimming club', which does not have very much to do with actual swimming. Pim Janssen and Jorik Waaijer, my most honored neighbors, for all the coffee breaks while working from home and the many, many evening walks to diverse ice cream shops all over Nijmegen. To you, I say: "Shantay, you *both* stay!" (RuPaul, once every season). Isabel Steenberg, for your 'judg-y', yet hilarious and honest conversations. You all taught me that, sometimes, the best way to move forward is to take a step back and take the day off - and thank Kevin.

To my oldest friends, Saartje Bekkers, Imke Bexkens, Manon Marinussen, Merel van der Aa, Laura Swinkels, Rick van de Loo and Ruud van Asseldonk, thank you for being there

for me from when I was an irresponsible teenager skipping classes, to when I became an irresponsible adult teaching classes. For our yearly celebration of ‘carnaval’ and ‘mintweedeschransdag’, because - after three times - *it’s tradition*. We have experienced many life phases together and I expect many more to come.

Anouk Boereboom, Anne Giesen, Gwen Reijnen and Rik Bosch, better known as ‘the goats’, who are always ready to help a woman in need (of a drink), thank you for all our nightly adventures. For your great taste in music, movies and fashion (some more than others). For being the weird and accepting group of people that you are. Cheers to many more festivals and camping trips!

Bas Berentsen and Saskia Uiterwaal, thank you for sharing your unique perspectives in philosophy and life since 2011. For the many craft beers we have conquered together and the many beautiful places upon which we have laid down our picnic blankets. Bas, my apologies for not wanting to play tennis with you in the rain. Or the cold. Or the heat. Maybe someday I will be a seasoned player like you, but I highly doubt it. Saskia, the word “taboo” really does not exist in your dictionary - and I am *all* for it.

Nina Janssen, who strongly believes that *Strange is Fun* (Klein Orkest, 1985), thank you for always being your genuine self, and for teaching me that every step in the right direction is a victory, no matter how small. For challenging my thoughts and proving me wrong sometimes (which can be hard to take, as a stubborn philosopher). For your patience with my many references to *Crazy Ex-Girlfriend* (2015) and *RuPaul’s Dragrace* (2009). Even though you had no clue what those references were about (it’s never too late to start watching), you always seemed to be on the same page as me. I really appreciated our walks through time and space and your excellent sense of humor, which often brought me the much-needed comic relief.

I want to thank my family. First of all, my parents, who always had my best interest at heart and who taught me to choose for what makes me happy, instead of what guarantees the most money or has the highest social status. In my case, this meant choosing to study philosophy, for which I have absolutely no regrets. Then, my brother Merlin, who is one of the funniest and most hard-working persons one can ever meet. We are definitely cut from the same cloth and I feel so lucky for knowing you since the beginning of times. Finally, to my

(unofficial) parents in law and grandma in law, thank you for all the 'gezelligheid' and for making me feel part of the family.

Last, but definitely not least, I want to thank Dilara. No matter the size of the tunnel, our car radios always seem to be tuned in on the same wavelength. My apologies for the cheesiness, but cheesiness and clichés only exist because many people agree something is true. Thank you for always understanding me, sometimes even before I do myself. For making even the objectively most boring vacation absolutely phenomenal. For eating cold pizza with me for breakfast with zero judgment. For always seeing the best in me, for your patience, your humor, for being my partner in crime and for always telling me the truth, beautiful or ugly. Because that is love (*The Favourite*, 2018). You are truly the most amazing person in my life.

About the author

Verna Jans was born on October 12, 1992 in Veghel and grew up in Boekel, the Netherlands. After obtaining her high school diploma at the Commanderie College Gemert in 2011, she enrolled in the Philosophy program at the Radboud University, Nijmegen. Here, she wrote her Bachelor thesis on the binary sex and gender difference in the context of Foucault's 'discourse'. After graduating in 2014, she started the Master program Philosophy of Science, Technology and Society at the University of Twente, Enschede, where she specialized in the Ethics of Technology. In 2016, half a year before graduating, she was granted the opportunity to start a PhD project at Maastricht University while still finishing her Master thesis on the ethics of creating human embryos for research on stem cell-derived gametes. Both endeavors turned out to be two sides of the same coin. As a PhD researcher, Verna investigated how new medically assisted reproductive techniques can be introduced into clinical practice in an ethically responsible way. Alongside her research, she tutored workshops concerning Health Ethics and Health Law to medical students. Currently, she works as a lecturer Ethics and Philosophy of Healthcare at IQ Healthcare, which is a scientific department focusing on quality improvement in healthcare, and as such forms part of the Radboud University Medical Center. Verna lives in Bottendaal, Nijmegen, with her partner Dilara and their cat Finn.

