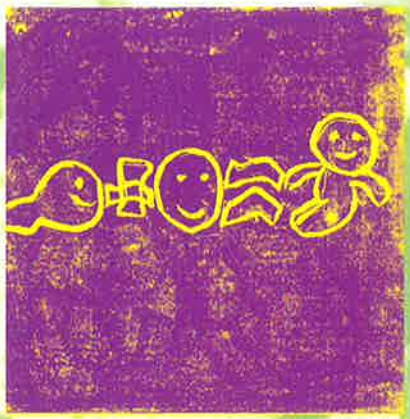
IVF pregnancies:
outcome and follow-up



Simone Buitendijk

IVF PREGNANCIES: OUTCOME AND FOLLOW-UP

Stellingen behorend bij het proefschrift

IVF pregnancies: outcome and follow-up

1. Eenlingzwangerschappen na IVF hebben slechtere uitkomsten dan die na natuurlijke conceptie (dit proefschrift).

2. Naast het vaker voorkomen van meerlingzwangerschappen na IVF en het ongunstiger profiel van IVF-moeders zoals hogere leeftijd, lijkt ook de IVF-procedure op zichzelf slechtere perinatale uitkomsten te veroorzaken (dit proefschrift).

3. Tweejarige IVF-kinderen hebben een iets grotere kans op een ontwikkelingsachterstand die vooral gerelateerd lijkt aan hun hoger risico op vroeggeboorte (dit proefschrift)

4. Vrouwen die op relatief jonge leeftijd een IVF-behandeling willen ondergaan, hebben vaker psychosociale problemen dan oudere IVF-kandidaten (dit proefschrift).

5. Doordat de overheid geen financiële middelen ter beschikking stelt voor langetermijnonderzoek naar kinderen geboren na IVF-zwangerschappen, wekt zij bij onvruchtbare paren de schijn dat de behandeling bewezen veilig is.

6. Koeien en vrouwen hebben genoeg met elkaar gemeen om de vergelijking tussen dracht, geboorte en ontwikkeling na IVP (In Vitro Productie) enerzijds en zwangerschap, geboorte en ontwikkeling na IVF anderzijds, wetenschappelijk waardevol te maken.

7. Zolang farmaceutische firma's niet door de overheid worden gedwongen om bij zwangerschappen na fertiliteitsbevorderende middelen, zinvolle en gedegen follow-up studies uit te voeren, is de verplichte postmarketing surveillance niet meer dan een geldverslindende papierwinkel zonder consequenties voor beleid.

8. Het is niet voldoende aangetoond dat IVF een effectieve behandelingsmethode is voor verminderde vruchtbaarheid.

9. Bij fertiliteitsproblemen geldt: 'With more patience, fewer patients'.

10. 'Iedereen kan toveren met een eitje'.
(titel afstudeerscriptie P.V.O. over de relatie tussen DES en IVF, Anita Direcks, 1985)

11. De internationale en nationale discussie over de thuisbevalling zou winnen aan wetenschappelijke zuiverheid wanneer de bewijslast voor veiligheid niet alleen bij verloskundigen ligt en de premisse niet langer geldt dat thuisbevallen riskant is tot het tegendeel is bewezen.

12. Pas wanneer Nederlandse mannen en vrouwen accepteren dat gedurende bepaalde periodes in het leven van een vrouw, haar borsten meer dan alleen een seksuele functie hebben, zal het bedroevend lage percentage borstgevoede pasgeborenen voldoende kunnen stijgen.

13. Het feit dat moderne managementboeken wat betreft de voorgestelde methoden veel overeenkomst vertonen met de huidige handboeken voor verantwoord ouderschap maakt het combineren van een leidinggevende functie met het grootbrengen van kinderen, iets gemakkelijker.

14. Het voordeel van het schrijven van een proefschrift is dat je het lot niet hoeft te laten bepalen wie eerste auteur wordt. (*n.a.v. Ten Geleide, pag. VII. Grondslagen der Epidemiologie. Vandenbroucke en Hofman of Hofman en Vandenbroucke 1993, Uitgeverij Bunge*).

15. Vrouwen kun je verdelen in twee groepen.
Maar ik zou het niet doen.
Het is ongelooflijk veel werk.
Finkers H. Een gynaecologisch praatje. Uit: het meisje met de eierstokjes 1997: 38-40. Novella Uitgeverij.

Leiden, 7 juni 2000
Simone E. Buitendijk

IVF PREGNANCIES: OUTCOME AND FOLLOW-UP

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus Dr. W.A. Wagenaar,
hoogleraar in de faculteit der Sociale Wetenschappen,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 7 juni 2000
te klokke 15.15 uur

door

Simone Elisabeth Buitendijk

geboren te Den Haag in 1958

PROMOTIECOMMISSIE

Promotores	:	Prof. dr. S.P. Verloove-Vanhorick Prof. dr. G.H.A. Visser (Universitair Medisch Centrum Utrecht)
Referent	:	Prof. dr. J.W. Wladimiroff (Erasmus Universiteit Rotterdam)
Overige leden	:	Prof. dr. E.V. van Hall Prof. dr. J.P. Vandenbroucke

'Oh, what a tangled web we weave, when first we practise to conceive!'
vrij naar: 'Oh, what a tangled web we weave when first we practise to deceive!'
(Sir Walter Scott, *Marmion*, canto 6, verse xvii, 1808)

Voor Caroline en Daniel

Cover
Design: Jaap van der Plas
Photography: Simone Buitendijk
Illustrations: Daniel en Caroline Figueroa, Jannie de Wit

This study has been carried out with financial support from the Dutch Ministry of Health, Welfare and Sports and from the Dutch Health Care Insurance Council (formerly known as Sick Fund Council). Their support is gratefully acknowledged.

CONTENTS

Chapter 1

Introduction

Introduction

The first Dutch IVF baby was born in 1983, in the Academic Hospital of Rotterdam. Five years earlier, the first ever IVF baby had been born in the UK. Most Western countries recorded the birth of their first IVF baby somewhere in the early eighties (Perone 1991). During the 1980's a large number of studies were conducted into the success of the IVF procedure (Edwards 1981, Wramsby 1981, Carson 1982, Fishel 1983, Jones 1983, Mahadevan 1983, Testart 1983, Edwards 1984, Quinn 1984, Naaktgeboren 1985). As the number of IVF procedures rose, concern was being expressed about the relatively low percentage of live births resulting from the procedure. In the United States, for instance, Congress enacted the 1992 Fertility Clinic Success Rate and Certification Act in response to consumer protection issues having been raised (Meikle 1999). By the early 1990's the course of IVF pregnancies and the subsequent development of the children had been much less extensively studied than the success rate of the procedure.

By then, some facts with respect to the course and outcome of IVF pregnancies had become apparent, however. It had become clear that IVF children more often are born prematurely and with low birthweight and that they are more likely to die in the perinatal period (Australian In Vitro Fertilisation Collaborative Group 1985, 1988, Nat Perinat Stat Unit 1988, MRC Working Party 1990, Beral 1990). Although these problems seemed primarily related to the higher risk of multiple pregnancies, it appeared they were also more frequent in singleton IVF pregnancies (Australian In Vitro Fertilisation Collaborative Group 1985, Cohen 1988, Saunders and Lancaster 1988, Beral 1990). It was yet unclear whether this increased risk in singleton IVF pregnancies was related to the IVF procedure itself or to the risk profile of the IVF mothers. Most IVF mothers were older and of lower parity and had had a longer period of attempting to become pregnant than women in the general population. Most studies carried out in those days did not employ the proper control groups to enable controlling for those factors. A number of the studies involved population figures as controls, or employed no control groups at all, thus rendering it impossible to adjust for factors such as maternal age and parity that are known to be different in the IVF population as compared to the general population of pregnant women. In an Australian cohort, the low birthweight rate in 700 singleton IVF children was reported to be 15.6% (Australian In Vitro Fertilization Collaborative Group 1988). A study of 741 IVF singletons found a low birthweight of 12%, not compared to a control group (Beral 1990). A study of 648 IVF singletons found 13% preterm deliveries compared to 6% in the general population, 11% versus 6% low birthweight and 17% small for gestational age (SGA) compared to 10% (Doyle 1992). The study did not control for confounders such as maternal age and parity. Olivennes (1993) compared 162 singleton IVF pregnancies with 263 pregnancies from stimulated cycles without IVF and with 5096 natural pregnancies. Prematurity occurred in 11.1% in the IVF group, in 6.1% in the stimulated group and in 4.4% in the spontaneous group. Low birthweight occurred in 11.1%, 6.5% and 3.6% and small for gestational age in 11.2%, 10.6% and 5.9%, respectively. The IVF mothers were older and more often primiparous than mothers in the two other groups. No attempt was made to control for these differences in the

analysis. Tan (1992) compared 763 IVF pregnancies to a control group matched for maternal age and parity and found a two times increased incidence of intrauterine growth retardation and preterm delivery in singleton IVF pregnancies.

In short, most of the studies into pregnancy outcomes after IVF singleton pregnancies carried out by the early nineties had shortcomings in the sense that they did not employ a formal control group or if one was used, the number of factors that was being controlled for was very limited. If controlling had taken place, only age and parity had been taken into account. It was clear that IVF singleton pregnancies have worse outcome than normally conceived singleton pregnancies, but it was much less clear which factors this was related to. IVF mothers are older on average and more often primiparous. They may be of higher socioeconomic status and different ethnicity and may have different risk taking behaviours (such as smoking and drinking) during pregnancy. They may suffer more from certain chronic conditions that may influence fertility as well as the outcome of pregnancy. The infertility itself, instead of the procedure, may cause worse pregnancy outcomes. To understand better the negative effects of the procedure on course and outcome of pregnancy, more demographic and pregnancy variables needed to be taken into account. Further, if after controlling for these potential confounders simultaneously, an effect of IVF remains, this would suggest a harmful effect of the procedure itself instead of a negative effect of maternal factors on the outcome. None of the studies carried out during the eighties and early nineties controlled for a large number of potential confounders at the same time, thus rendering it impossible to draw a firm conclusion.

In a few studies, attempts had been made to control for the infertility as an indication for the IVF treatment. In one of these studies, couples for whom tubal obstruction was the main indication for the IVF treatment, had lower reported risks of adverse perinatal outcomes such as low birthweight (Doyle 1992) than did couples who underwent IVF for reasons related to idiopathic or male infertility. Another study (Saunders 1988) compared spontaneous pregnancies to IVF pregnancies and to pregnancies in untreated women on the waiting list for IVF, finding that the rate of prematurity was comparable in the latter two groups and that both were higher than in the first group. These data suggest that infertility of male or unknown causes may be a risk factor in itself. Olivennes (1993) found higher risks of prematurity and low birthweight in the IVF singleton group than in the stimulated group, but still lower rates in the group of spontaneously conceived pregnancies, indicating there may be a risk attached to the ovulation stimulation and a separate risk attached to the IVF procedure itself.

Another area of interest were congenital defects in IVF children. The total number of congenital defects reported in children in national IVF studies carried out by that time did not show an increase (Nat Perinat Stat Unit 1988, MRC Working Party Med Research Int 1990). One study, however, showed an increase in number of neural tube defects (Med Research Int 1991), a defect that has been linked to Clomifene (Mills 1990) which was commonly used in IVF treatments in those days. Further, very few and only relatively small

studies had been carried out into the development of the children born after IVF. It was therefore not clear whether IVF children developed normally. Another area of interest were the problems that women treated with IVF may encounter. It was known at that time that superovulation as employed in the IVF procedure may lead to Ovarian Hyperstimulation Syndrome (OHS), but it was less clear whether long-term effects may appear. A number of authors had speculated that since suppression of ovulation by oral contraceptive use decreased the risk of ovarian cancer, (repeated) superovulation may increase the risk (Fathalla 1971, Casssagrande 1979, van Hall 1988, Fishel 1989). No formal studies had been carried out that could prove or disprove this theory. Last, the psychological effects of undergoing IVF treatment had not been studied in a structured way.

In the early nineties it was obvious from the relative paucity of data that need existed for a large scale, well designed study. In 1991, a first meeting took place with IVF clinicians and embryologists from all licensed IVF centres to discuss a possible national collaborative study to elucidate the questions that remained with respect to the potential effect of IVF on the course of the subsequent pregnancies and development of the children. The meetings resulted in a joint proposal in 1992. The study questions in our research proposal were phrased as follows:

1. Do IVF multiple and singleton pregnancies have different risks of preterm birth, birthweight, perinatal and neonatal morbidity and congenital defects than non-IVF pregnancies do?
2. If they do, is the different risk related to the profile of the IVF mothers and/or the indication for treatment?
3. Do pregnancies after cryopreservation have different outcomes than those after fresh embryo transfer?
4. Does the psychomotor development of IVF children at the age of two differ from that in a control population?
5. What is the quality of life of couples before and 1.5 years after the start of IVF treatment?
6. Do women who were treated with IVF more often than other women experience gynaecological complaints?

It was obvious from the onset that if the study would show a negative effect of IVF on these outcomes, after correction for maternal and pregnancy variables, it would not be possible given the research design, to pinpoint specific aspects of the IVF procedures as the main causative factors. It was felt, however, that first and foremost the question whether IVF pregnancies were more likely to be related to negative outcomes, with or without taking into account maternal and pregnancy variables, needed to be answered.

Methodology of the study

The proposed study duration was 5 years. During 1993, preparations were made to start the actual study. Informed consent forms were developed, the study protocol was discussed in each participating clinic separately and over 20 medical ethics committees studied and

approved the proposed methods. This resulted in a general approach with individualised forms and procedures for each participating centre. It was decided by the participating centres and the researchers that informed consent should be asked after a pregnancy had been established in order to avoid asking consent from a large number of women who would not be able to achieve pregnancy. Since our main hypotheses concerned outcome of pregnancy and not the effectiveness of the procedure, this was judged to be the most practical approach. January 1994, women who had become pregnant after IVF were first asked to participate. It was estimated that approximately 800 singleton IVF births were needed to test our primary hypotheses and that data collection should last two years i.e. until the end of 1995. It was planned to test 200 IVF children at the age of two and to study around 500 IVF couples with respect to their psychological status.

Two conrol groups were employed. First, pregnancies from the Social Medical Survey of Children attending Child Health Clinics (SMOCC) served as controls (Herngreen 1992). In the SMOCC-study, data were collected through postpartum interview with all mothers from liveborn infants in the areas of 21 different child health clinics. Of 2151 babies born between April 1st 1988 and October 31st 1989, data were gathered concerning maternal, socio-demographic and health characteristics, the course of the pregnancy and the health of the newborn, combined with data routinely collected by the Child health Clinics on course and outcome of pregnancy. The routinely collected data were obtained from the midwife or physician who had provided care during pregnancy and/or delivery. If a pediatrician had been consulted, information from that source was added to the database. Since the SMOCC-dataset is restricted to live-born babies only, a second control dataset was employed that also contains information on infants that died in the perinatal period. In order to create this second control group representative of all births in the Netherlands, the 1995 data of two Dutch national databases were combined: the National Perinatal Database of obstetrical care delivered by midwives and general practitioners (LVR-1) and the National Perinatal Database of obstetrical care delivered by obstetricians (LVR-2). Midwives, GP's and obstetricians register information about pregnancy, delivery and puerperium of pregnancies with a gestational age of at least 16 completed weeks, in most cases during the first week after delivery. In 1995, 89% of all midwives and obstetricians participated in the LVR. The GP's participation was very limited. Records in the joint database do not have a unique number. Further, pregnancies (and children) can occur twice in the joint database in case women are referred during pregnancy from a midwife to an obstetrician. An aggregated perinatal data file representative of all 1995 births in the Netherlands was created. In order to identify each anonymous child as a single record, for each record a unique key was defined. A number of identifying variables were used such as mother's postal code and child's date of birth. This dataset has been used in a study on perinatal deaths in the Netherlands and has been validated by comparing the number of deaths in this database to the number in the 1995 perinatal database of the Dutch Central Bureau of Statistics (CBS). Methods are described in detail elsewhere (Anthony, in preparation). As a last step, an attempt was made to identify IVF pregnancies in the LVR-dataset.

A number of key variables (among which mother's birth date, birth date of the child, sex of the child) were used to identify the IVF pregnancies in the LVR-dataset. Also, a variable in the LVR-dataset indicating whether IVF treatment had taken place, was used to identify IVF pregnancies. Those pregnancies that likely were IVF pregnancies, were subsequently removed from the LVR-control dataset.

The study was carried out by TNO Prevention and Health in collaboration with the IVF Working Group of the Dutch Society of Obstetricians and Gynaecologists. The Netherlands has 12 IVF centres holding laboratory facilities, of which 8 are academic and 4 non-academic. In these IVF centres, the entire IVF procedure can be performed, from the (patient) intake to the embryo transfer and the follow-up. The IVF centres collaborate with so-called transport hospitals. In the transport hospitals, all parts of the IVF procedure are being performed with the exception of the laboratory phase and the embryo transfer. The laboratory phase is being carried out in one of the IVF centres. Some transport hospitals co-operate with other hospitals called satellite hospitals (Roest et al., 1995). In this construction, three hospitals participate in the treatment procedure. The harvesting of the oocytes takes place in the transport hospital, the laboratory phase and embryo transfer are being performed in the IVF centre and the other parts of the procedure i.e., patient selection, hyperstimulation, cycle monitoring and follow-up after embryo transfer are being carried out in the satellite hospital. In the Dutch IVF program in 1994, 12 IVF centres, 23 transport hospitals and 6 satellite hospitals took part. All hospitals except one transport hospital agreed to participate in our study.

From January 1st 1994, to December 31st 1995, women who were interested in participating were asked written informed consent. It was obtained shortly after the IVF treatment had shown to have resulted in an ongoing pregnancy, diagnosed primarily with fetal heart action at ultrasound. Women who subsequently miscarried before 16 weeks of pregnancy, were excluded from the study. Data on infertility history and treatment schedule were obtained through the participating IVF clinic. Two months after the expected date of delivery the IVF patient received a mailed questionnaire containing questions about her obstetric and medical history as well as the course and outcome of her pregnancy. The provider of the prenatal care (general practitioner, midwife or obstetrician) was asked to provide information concerning the prenatal period and delivery, through a questionnaire. If the newborn had been in paediatric care a copy of the paediatrician's report was obtained. The IVF clinic that had provided the treatment was asked to provide information on infertility history as well as on characteristics of the IVF treatment.

The study was completed (apart from the investigation of the IVF mothers' gynaecological complaints) at the end of 1999. A total of 2043 pregnant IVF women took part, who had 2636 children, 1502 of them being singletons. It had turned out that the use IVF had increased quite dramatically between 1992 and 1994/95 as a result of which we were able to collect data on more pregnancies and deliveries than we had originally planned. Of the children a subgroup of 200 were tested for their psychomotor development. Close to 450

couples filled out questionnaires measuring their psychological status before the IVF treatment and 1.5 years later. Close to 40 IVF hospitals took part in the data collection. All 12 licensed clinics participated as well as the majority of hospitals collaborating with them.

More details about data collection methods and choice of control groups are provided in this manuscript. The answers to research questions 1,2,3 and 4 as outlined above will be described in this manuscript as well as part of question 5 (the psychological status of IVF couples before the IVF treatment). A number of publications are in preparation. They will appear separate from this manuscript. These relate to:

1. outcome of IVF twin pregnancies
2. congenital defects in IVF children
3. quality of life of IVF couples 1.5 years after the start of the treatment
4. relationship between levels of stress and chance of success in IVF couples
5. gynaecological complaints of IVF mothers

Aims of the thesis

The aims of this thesis are to answer the following research questions: do IVF multiple and singleton pregnancies have different risks of preterm birth, growth retardation, birthweight and perinatal mortality than non-IVF pregnancies do? If they do, is the different risk related to the profile of the IVF mothers and/or the indication for treatment? Do pregnancies after cryopreservation have different outcomes than those after fresh embryo transfer? Does psychomotor development of IVF children at the age of two differ from that in a control population? What is the quality of life of couples before the start of IVF treatment?

Outline of the thesis:

- Chapter 2 provides an overview of the literature;
- Chapter 3 describes the risk of preterm birth in IVF children as compared to controls;
- Chapter 4 describes the risk of low birthweight and growth retardation in IVF singletons as compared to controls;
- Chapter 5 describes the outcomes of pregnancies after cryopreservation as compared to fresh embryo transfer;
- Chapter 6 describes psychomotor development of IVF children at the age of two compared to development of controls;
- Chapter 7 describes couples' estimated chance of success;
- Chapter 8 describes health-related quality of life of IVF couples planning to undergo IVF;
- Chapter 9 provides reflective conclusions.

References

Anthony S, van der Pal KM, Graafmans WC et al. The reliability of perinatal and neonatal mortality rates: differential underreporting in linked professional registers versus Dutch national statistics. Submitted to the American J of Public Health.

Australian In-Vitro Fertilization Collaborative Group. High incidence of preterm births and early losses in pregnancy after in vitro fertilization. BMJ (Clinical Research ed.) 1985;291:1160-3.

Australian In-Vitro Fertilization Collaborative Group. In-vitro fertilization pregnancies in Australia and New Zealand, 1979-1985. Med J of Australia 1988;148:432-46.

Beral V, Doyle P, Tan SL, Mason BA, Campbell S. Outcome of pregnancies resulting from assisted conception. British Medical Bulletin 1990;46:753-68.

Carson RS, Trounson AO, Findlay JK. Successful fertilisation of human oocytes in vitro: concentration of estradiol-17 beta, progesterone and androstenedione in the antral fluid of donor follicles. J Clin Endocrinol Metab 1982;55(4):798-800.

Cassagrande JF, Louie EW, Pike MC, Roy S, Ross RK, Henderson BE. Incessant ovulation and ovarian cancer. Lancet 1979;2:170-2.

Cohen J, Mayaux MJ, Guihard-Moscaro L. Pregnancy outcomes after in vitro fertilisation. A collaborative study on 2342 pregnancies. Annals of the New York Academy of Science, 1988;541;1-6.

Doyle P, Beral V & Maconochie N. Preterm delivery, low birthweight and small-for gestational-age in liveborn singleton babies resulting form in-vitro fertilization. Human Reproduction 1992;7:425-8.

Edwards RG. Test-tube babies, 1981. Nature 1981;293(5830):253-6.

Edwards FG, Fishel SB, Cohen J, Fehilly CB, Purdy JM, Slater JM, Steptoe PC, Webster JM. Factors influencing the success of in vitro fertilization for alleviating human infertility. J In Vitro Fert Embryo Transf 1984;1(1):3-23.

Fathalla MF. Incessant ovulation, a factor in ovarian neoplasia? Lancet 1971;2:163.

Fishel SB, Edwards RG, Walters DE. Follicular steroids as a prognosticator of successful fertilization of human oocytes in vitro. J Endocrinol 1983;99(2):335-44.

Fishel S, Jackson P. Follicle stimulation for high tech pregnancies: are we playing it safe? Br Med J 1989;299:309-11.

Hall van, EV. Over manipulatie van de menselijke voortplanting. Medisch Contact 1986;46:1479-82.

Jones HW Jr, Acosta A, Andrews MC, Garcia JE, Jones GS, Mantzavinos T, McDowell J, Sandow B, Veeck L, Whibley T, Wilke C, Wright G. The importance of the follicular phase to success and failure in in vitro fertilization. Fertil Steril 1983;40(3):317-21.

Mahadevan MM, Trounson AO, Leeton JF. The relationship of tubal blockage, infertility of unknown cause, suspected male infertility, and endometriosis to success of in vitro fertilization and embryo transfer. Fertil Steril 1983;40(6):755-62.

Medical Research International and the Society of Assisted Reproductive Technology. The American Fertility Society. In vitro fertilisation/embryo transfer in the United States: 1989 results from the National IVF-ET Registry. Fertil Steril 1991;55:14-23.

Meikle SF, Danel I, Wilcox LS. Surveillance of assisted reproductive technology in the United States. Int J of Technology Assessment in Health Care 1999;15:11-4.

Mills JL, Simpson JL, Rhoads GG, et al. Risk of neural tube defects to maternal fertility and fertility drug use, Lancet 1990;2:103-4.

MRC Working Party on Children Conceived by In Vitro Fertilization. Births in Great Britain resulting from assisted conception 1978-87. BMJ (Clinical Research ed.),1990;300:1229-33.

Naaktgeboren N, Devroey P, Traey E, Wisanto A, Van Steirteghem AC. Success of in vitro

fertilization and embryo transfer in relation to the causes of infertility. Acta Eur Fertil 1985;16(4):281-7.

National Perinatal Statistics Unit and Fertility Society of Australia. IVF and GIFT pregnancies: Australia and New Zealand, 1987. Sydney National Perinatal Statistics Unit, 1988.

Perone N. In vitro fertilization and embryo transfer: A historical perspective. J of Reproductive Medicine 1994;39:695-700.

Quinn P, Warnes GM, Kerin JF, Kirby C. Culture factors in relation to the success of human in vitro fertilization and embryo transfer. Fertil Steril 1984;41(2):202-9.

Roest J, Verhoef AA, Lent M van, Huisman GJ, Zeilmaker GH. Results of decentralized in vitro fertilization treatment with transport and satellite clinics. Hum Reprod 1995;10:563-7.

Saunders DM, Mathew M, Lancaster PAL. The Australian Register: current research and future role: a preliminary report. Annals of the New York Academy of Science, 1988;160:209-13.

Testart J, Lassalle B, Frydman R, Belaisch JC. A study of factors affecting the success of human fertilization in vitro. II. Influence of semen quality and oocyte maturity on fertilization and cleavage. Biol Reprod 1983;28(2):425-31.

Van Hall EV. Manipulation of human reproduction. J Psychosom Obstet Gynaecol 1988;9:207-13.

Wilcox LS. Federal oversight of assisted reproduction: Public Health, consumer protection, and public resources. Women's Health Issues 1996;6:146-7.

Wramsby H, Kullander S, Liedholm P, Rannevik G, Sundstrom P, Thorell J. The success rate of in vitro fertilization of human oocytes in relation to the concentrations of different hormones in follicular fluid and peripheral plasma. Fertil Steril 1981;36(4):448-54.

Chapter 2

Children after in vitro fertilization: An overview of the literature

S.E. Buitendijk

Buitendijk SE. Children after in vitro fertilization. Int J of Technology Assessment in Health Care 1999;15:52-65.

Abstract

This article provides an overview of the effects of the IVF procedure on the children born from it. One of the main problems with IVF to date remains the high incidence of multiple pregnancies, which carry an inherent higher risk of preterm delivery and, therefore, of increased morbidity and mortality in the newborns. Further, IVF singleton pregnancies and IVF twin pregnancies compared to control singleton or twin pregnancies appear to be at higher risk of preterm birth and low birthweight. Whether this is an effect of the procedure per se, or is related to maternal factors, or a combination of both, remains to be studied. The risk of congenital malformations does, with the available data, not seem to be elevated. As of now, it remains unclear whether embryo freezing is a safe procedure. Psychomotor development of IVF children does not seem to be disturbed, possibly because of the high quality of IVF parenting. However, until further and more extensive studies are conducted, it remains unclear whether IVF poses long term risks for the children.

Key words: IVF, side-effects, follow-up studies, review, IVF pregnancies, IVF children

Introduction

In Vitro Fertilisation (IVF) has in most Western countries been introduced during the early 1980's as a new infertility treatment. It's introduction was not coupled with formal evaluation of the possible adverse effects on the health of the women, the course of the pregnancies or the health of the babies and children born from it. IVF has nevertheless become widespread during subsequent years. In the Netherlands, the first IVF baby was born in 1983. In 1997, approximately 1% of all Dutch babies were born after IVF. The development of IVF has led to the introduction of related new techniques, such as ICSI (IntraCytoplasmatic Sperm Injection). ICSI can help overcome severe male factor infertility. Instead of bringing together a number of egg cells with millions of sperm in a petrie dish (which is the case with IVF), with ICSI, each egg cell is being injected with a single sperm cell. Even in 1998, many of the questions concerning short- and long-term effects of IVF on the children, have not been sufficiently answered. Very much like IVF, spin-off techniques such as ICSI have become popular without careful evaluation of potential side effects.

Each of the different phases of the IVF treatment may carry in them the risk of side-effects for the developing embryo. First, hormonal treatment of the woman takes place, to enable ripening of more than one follicle and, often, to better 'time' the optimal moment of follicle aspiration. Most of the hormonal preparations used have not been studied sufficiently for potential teratogenic effects or for effects on the course of a subsequent pregnancy. Second, oocyte retrieval takes place. Theoretically, this phase could carry the risk of mechanical damage to the egg cells. Third, eggs and pre-treated sperm are brought together in a petrie dish to enable fertilisation. Fertilised eggs are being transferred into a special culture medium for further growth.

It is as yet unclear whether substances in the culture medium may have teratogenic effects on the growing embryo. Fourth, if dividing embryos have developed, they can be transferred back into the uterus. Potentially, the mechanical process of placing the embryo into the uterus, may lead to a less optimal place of nidation and placentation than would be the case in the natural situation. Fifth, surplus embryos may be frozen to be transferred in a later cycle. Freezing may potentially be harmful to the embryo.

Since the late seventies, a vast number of articles have been published concerning the likelihood of pregnancy per IVF treatment cycle. Patient characteristics, but even more, aspects of the treatment protocol have been studied extensively for their effect on the succes rate. The course of subsequent pregnancies and, especially, the development of children born after IVF, has been much less extensively studied. In this paper, an overview will be presented from the published research data concerning the outcome of IVF pregnancies and the subsequent development of the children. Risks for the woman will not be discussed.

Multiple pregnancies

Since the success per cycle of IVF depends to a great extent on the number of embryos being placed in the uterus, in the majority of IVF treatments, more than one embryo is transferred. Consequently, a large percentage of IVF pregnancies are multiple pregnanies. In most publications, mention is made of a multiple pregnancy rate of 25% to 30%. Approximately 3% of IVF pregnancies consist of triplet or higher order pregnancies (American Society for Reproductive Medicine 1995, Dulioust 1995, FIVNAT 1995). Thus, almost half of all children born after IVF are part of a multiple pregnancy. Multiple pregnancies inherently carry a higher risk of complications, of which prematurity, intrauterine growth retardation and perinatal mortality are the most important (Botting 1987, Kiely 1992). In most IVF cohorts an excess of prematurity, low birthweight and perinatal mortality is observed, which seems primarily related to the high incidence of multiple pregnancies. A large French study of IVF pregnancies (FIVNAT 1995) showed a prematurity rate of 29%, a low birthweight rate of 36% and a perinatal mortality rate of 27 per thousand, the excess in adverse outcomes mostly caused by the multiple pregnancy rate. D'Souza (1997) reporting outcomes of an IVF cohort, indicated that the higher percentage of preterm deliveries was largely due to the multiple births and that they contributed to adverse neonatal conditions in 45% of all IVF children in the cohort. These data are in accordance with figures on outcomes of multiple pregnancies in other IVF cohorts.

The rate of triplet and higher order pregnancies largely depends on the clinic's policy with respect to number of embryos being transferred. In the Netherlands, the present policy is not to transfer more than two embryos, unless their individual quality is very low and/or the woman is of higher age, in which case the expected chance of success per embryo is low. This policy is expected to lead to a reduced rate of multiple pregnancies in the near future.

Low birthweight, preterm birth and small for gestational age in IVF singletons

Many authors have reported an increased incidence of adverse pregnancy outcomes in IVF singleton children. A number of these studies involved population figures as controls, or employed no control groups at all, thus rendering it impossible to adjust for factors such as maternal age and parity, that are known to be different in the IVF population as compared to the general population of pregnant women. In an Australian cohort, the low birthweight rate in 700 singleton IVF children was reported to be 15,6% (Australian In Vitro Fertilization Collaborative Group 1988). A study of 741 IVF singletons found a low birthweight rate of 12%, not compared to a control group (Beral 1990). A study of 648 IVF singletons found 13% preterm deliveries compared to 6% in the general population, 11% versus 6% low birthweight and 17% small for gestational age (SGA) compared to 10% (Doyle 1992). The study did not control for confounders such as maternal age and parity. Olivennes (1993) compared 162 singleton IVF pregnancies with 263 pregnancies from stimulated cycles without IVF and with 5096 natural pregnancies. Prematurity occurred in 11.1% in the IVF group, in 6.1% in the stimulated group and in 4.4% in the spontaneous group. Low birthweight occurred in 11.1%, 6.5% and 3.6% and small for gestational age in 11.2%, 10.6% and 5.9%, respectively. Only the differences between the IVF group and the spontaneous group reached statistical significance. The IVF mothers were older and more often primiparous than mothers in the two other groups. No attempt was made to control for these differences in the analysis. Rufat (1994) compared IVF singletons to population controls (not adjusting for possible differences between the IVF group and the control group) and found a risk of prematurity of 12% compared to 5.6% and a 15% birthweight under the 10th percentile. A large national French study (FIVNAT 1995) found a prematurity rate of 9.4% and a low birthweight rate of 11.2% in IVF singletons in almost 4000 IVF singletons. These results too, were not compared to those in a formal control group.

In a number of studies into outcomes of IVF singleton pregnancies, attempts have been made to control for certain factors. Tan (1992) compared 763 IVF pregnancies to a control group matched for maternal age and parity and found a two times increased incidence of intrauterine growth retardation and preterm delivery in singleton IVF pregnancies. A study of 140 singleton IVF pregnancies and an equal number of control pregnancies, matched for maternal age, height, weight and parity found that IVF pregnancies were one week shorter in duration and that the children weighed 175 grammes less than controls (Verlaenen 1995). Petersen (1995) in a small study compared 70 IVF singleton-pregnancies to 70 pregnancies in women with fertility problems, of comparable age and parity, who had not undergone IVF. The preterm delivery rate was 7% in the singleton IVF group, compared to 4% in the singleton control group, the low birthweight rates were 16% and 0%, respectively. A study of 150 singleton IVF children, comparing them to an equal number of controls born at term and matched for sex of the child and social class, found that IVF singletons were on average born one week earlier than the controls and weighed 400 grammes less (Dulioust 1995). Reubinoff et al. (1997) in a study of 260 IVF singletons compared to 260 controls, found no difference in percentage of low birthweight and smal for gestational age, after pairwise matching for maternal age, parity, ethnic origin and

location and place of birth. The IVF children were more premature, a difference which largely seemed due to the increase in caesarean sections before 37 weeks in the IVF group. A number of other studies comparing IVF singletons to other singletons also found that IVF singletons have shorter gestational age and low birthweight (MRC Working Party on Children Conceived by In Vitro Fertilization 1990, Wada 1994, Gissler 1995, Saunders 1996). A summary of these data is being presented in table 1.
Most of the studies into pregnancy outcomes after IVF singleton pregnancies, have shortcomings in the sense that they do not employ a formal control group or if one is used, the number of factors that is being controlled for is limited. The evidence, however, that IVF singleton pregnancies are more likely to end preterm and result in low birthweight babies, is overwhelming. It is clear from the above mentioned literature that IVF singleton pregnancies have worse outcomes than normally conceived singleton pregnancies, although it is less clear whether this is related to the IVF procedure per se, or to other factors. IVF mothers are older on average and more often primiparous. They may be of higher socioeconomic status and may have different risk taking behaviours (such as smoking and drinking) during pregnancy. They may suffer more from certain chronic conditions that may influence fertility as well as the outcome of pregnancy. The infertility itself, instead of the procedure, may cause worse pregnancy outcomes. If after controlling for these potential confounders simultaneously, an effect of IVF remains, this would suggest a harmful effect of the procedure itself instead of a negative effect of maternal factors on the outcome. Unfortunately, none of the studies carried out so far have controlled for a large number of potential confounders at the same time, thus rendering it impossible to draw a firm conclusion as yet.

In a few studies, attempts have been made to control for the infertility as an indication for the IVF treatment. In one of these studies, couples for whom tubal obstruction was the main indication for the IVF treament, had lower reported risks of adverse perinatal outcomes such as low birthweight (Doyle 1992) than did couples who underwent IVF for reasons related to idiopathic or male infertility. Another study (Saunders 1988) compared spontaneous pregnancies to IVF pregnancies and to pregnancies in untreated women on the waiting list for IVF, finding that the rate of prematurity was comparable in the latter two groups and that both were higher than in the first group. These data suggest that infertility of male or unknown causes may be a risk factor in itself. Olivennes (1993) found higher risks of prematurity and low birthweight in the IVF singleton group than in the stimulated group, but still lower rates in the group of spontaneously conceived pregnancies, indicating there may be a risk attached to the ovulation stimulation and a separate risk attached to the IVF procedure itself.

If the IVF procedure itself is a contributing factor to the higher rate of low birthweight and/or the preterm labour in IVF singleton pregnancies, instead of the mothers' profile, a number of possible explanations can be offered. Some studies have linked superovulation to reduced birthweight and growth retardation in humans. Elevation of insulin-like growth factor binding protein (IGFBP-1) is the suggested cause (Lino 1986, Howell 1989, Wang 1991, Johnson 1993, Johnson 1995), which would imply that the ovulation induction within the IVF procedure may be harmful to foetal development. Another study suggests

that a higher incidence of abnormal placental shapes exists in IVF singleton pregnancies (22% vs 6% in the control group) and that abnormal umbilical cord insertions are more often found (Jauniaux 1990). The phenomenon may be related to inadequate orientation of the blastocyst after IVF embryo transfer which may cause inferior placental functioning. In cattle and sheep assisted reproduction, culture, embryo micromanipulation and transfer, are known to lead to larger sized offspring, which in these species is a sign of pathology (Wennerholm 1998, Walker 1992). In animal assisted reproduction, the animals donating the gametes, nor the recipients, suffer from infertility. Therefore, these adverse effects, if real, have to be related to aspects of the procedure itself.

IVF multiple pregnancies compared to other multiple pregancies

IVF multiple pregnancies may be at higher risk of complications than multiple pregnancies that have been conceived naturally. A number of studies comparing naturally conceived multiple pregnancies to IVF multiple pregnancies either combine all multiple pregnancies in their analysis, or suffer from numbers too low to enable firm conclusions (Rizk 1991, Friedler 1992, Tan 1992).

Two recent studies, however, have specifically compared IVF twin and triplet pregnancies, to twin and triplet pregnancies conceived otherwise (Bernasko 1997, Friedler 1992). One study comparing 105 IVF twins to 297 twins after natural fertilisation (Bernasko 1997), found a statistically significantly higher incidence of low birthweight (72% vs. 59%) and discordant birthweight (23% vs. 14%) in the IVF twin pregnancies compared to control twin pregnancies, after controlling for maternal age and parity. Prematurity rates did not differ between the two groups. Discordant fetal growth is a complication that especially occurs in monochorionic pregnancies, probably as a result of an imbalance in the blood flow through the placental arteries. It is less likely to occur in dizygotic twins, who have separate chorionic membranes. Therefore, one might expect that discordant birthweight is a phenomenon more likely to be seen in spontaneous twin pregnancies (25% of which are monozygotic), than in IVF pregnancies, that are essentially made up of dizygotic pregnancies only. The authors offer the possible explanation that IVF twin pregnancies are more often of unlike sex than the control twin pregnancies. Unfortunately, no analysis was carried out comparing same sex twins in both cohorts separately. Both the higher rate of discordant birthweight and higher rate of the low birthweight in the IVF group may be related to decreased functioning of the placenta, as also may be the case in IVF singleton pregnancies. The other study (Friedler 1992) compared 56 triplet pregnancies after 'artificial reproductive technology' to 82 triplet pregnancies after ovulation induction and 13 triplet pregnancies conceived spontaneously. Mean gestational age of triplets following ART (33.2 weeks) was not significantly different from those conceived following gonadotrophin stimulation (33.4 weeks) or stimulation with clomiphene citrate (34.2 weeks), but was significantly shorter than in the pregnancies after spontaneous conception (35.3 weeks). Mean foetal birthweight was 1743 grammes, 1683 grammes, 1863 grammes and 1963 grammes respectively. Only the difference between the birthweights in the ART and in the spontaneous group was statistically significant. No difference was found in low birthweight rates between the groups, although the very low birthweight rates (< 1500

grammes) following ART and ovulation induction were significantly different from those in the spontaneous group (31%, 30% and 10%, respectively).

Although the groups compared are still small in terms of absolute numbers, these results seem to indicate that a possible adverse effect from ART (and therefore from IVF) is more likely to stem from the ovulation stimulation than from other aspects of the procedure. The analysis did not control for maternal characteristics. It is not possible, therefore, to conclude whether age, parity or other maternal characteristics related to infertility may have (partially) determined the worse outcome in the ART and stimulation group compared to the group of naturally conceived pregnancies. Another recent study, in contrast, comparing a group of 72 IVF twin pregnancies, 82 twin pregnancies after ovulation induction and 164 spontaneous twin pregnancies did not find a difference in prematurity rate (39%, 45% and 40%), SGA (18%, 23% and 23%) and perinatal mortality (3%, 3% and 4%) between the three groups (Olivennes 1996).

Child behaviour

Several investigators have studied the behaviour and the cognitive and psychomotor development of IVF children. One of the first studies (Mushin 1986) investigated 33 IVF children with the Bayley Scales of Infant development, between the ages of 12 and 20 months, comparing them with a norm population. The IVF children performed within the normal range and the subgroup with developmental problems consisted mainly of IVF children that were born prematurely. Morin (1986) studied 83 IVF children and compared them to 93 controls matched for age of the infant, multiple conceptions, sex, race and maternal age. No differences were found between the IVF and the control group. The largest so far (Brandes 1992) studied 116 IVF children at 12 to 45 months of age and compared them to a control group matched for gestational age, multiple pregnancy and a number of maternal characteristics, using the Bayley Scales and the Stanford Binet Scales. The development of the IVF children did not differ from that of the control group. A small study (Raoul-Duval 1994) compared 33 children born after IVF to 33 children born after ovulation induction without IVF and to 33 naturally conceived children at the age of three years. Only term singletons were studied. No differences in development were found between the three groups. Another small study (Ron-El 1994) looked at 30 children born after IVF, who were at least two and a half years of age, and compared them to a singleton control group.
Recently, a study compared 99 IVF children aged 33 to 85 months, to a population control group in terms of their development (measured with the Griffith Scale) and their behaviour (measured with the Child Behaviour Check List) (Cederblad 1996). 34% Of the IVF children were part of a multiple birth and 28% were born prematurely. Although both the IVF group as a whole and the preterm group had scores comparable the norm population group, IVF children who had experienced normal conditions did have higher developmental quotients on the Griffith Scale than those who had been born prematurely. These studies have been summarized in table 2.

Most studies into child development after IVF carried out up till now comprise relatively small numbers of children and have a limited period of follow-up. Many disorders, such as school performance problems and attention deficit disorders, can only be diagnosed at older ages. Still, the figures published so far, do not seem to give rise to great concerns. One possible explanation of why IVF children may be developing well, is that the quality of parenting of IVF parents is superior to that of control parents, thus overcoming the possible adverse effects of the IVF treatment per se, or of the prematurity related to it. One study compared 45 families with a child conceived by IVF, with 45 families with a child conceived by donor insemination, 45 families with a naturally conceived child and 55 families with an adopted child (Golombok 1996). The families were matched for social class of the parents and age and sex of the child. The children were between the ages of 4 and 8 years. No differences were found between the groups in terms of emotional or behavioural variables, but parent-child interaction was superior in the IVF families, possibly indicating a higher degree of 'wantedness'. This difference in interaction patterns may well influence the IVF childrens' score on developmental tests. A study looking at behaviour towards the infant of 65 primiparous IVF mothers four months post partum, however, found that IVF mothers, more often than the controls, reported lower self esteem and lower maternal self efficacy (MacMahon 1997). Mothers who had received more than one treatment cycle received lower scores. It is not clear whether the mothers' perception of themselves may persist until later in the IVF childrens' lives.

Birth defects

A number of authors have studied the occurrence of birth defects in children born after IVF (Trounson 1984, Seppala 1985, Australian In Vitro Fertilization Collaborative Group 1985, Australian In Vitro Fertilization Collaborative Group 1988, Morin 1989, Cohen 1988, MRC Working Party on Children Conceived by In Vitro Fertilization 1990, Rizk 1991, American Society for Reproductive Medicine 1995). The number of newborns studied ranged from fewer than a 100 to over 3,000. None of these studies found an increased risk of birth defects compared to the general population estimates, for all defects combined, nor for specific groups of defects. Only one study (Lancaster 1987) mentions a statistically significant increase in two types of birth defects, namely a higher prevalence of transposition of the great vessels and of spina bifida in IVF children than in the general population. This report has not yet been confirmed in other studies. Theoretically, birth defects after IVF may be increased in incidence because of the induction of chromosomal abberations, an increase in fertilisation rate by abnormal spermatozoa, or the actions of physical and chemical teratogens (Biggers 1981). At the moment there is no evidence of such effects. However, to find a significant increase in groups of birth defects, or in isolated defects, instead of in all defects combined, larger numbers of cases are often needed than are part of the studies here mentioned. In order to test specific hypotheses concerning IVF and well defined birth defects, in the future pooled analyses of some of the larger studies may provide the necessary numbers.

Effects of cryopreservation

The superovulation method as it is employed in most IVF treatments, often yields a high number of oocytes to be fertilised. If too many embryos are thus being created, supernumary embryos can be frozen to be thawed in subsequent cycles. The first successful pregnancy after cryopreservation was decribed in 1983 (Trounson). Cryopreservation and thawing involve major cellular changes and it is not clear whether they may have adverse effects for the offspring. A number of studies into the perinatal outcome of pregnancies after cryopreservation have been reported (Reubinoff 1997, Sutcliffe 1995, Olivennes 1997, Wada 1994, Wennerholm 1997, Wennerholm 1998). Wada (1994) did not find any chromosomal abnormalities in the cryopreserved group and found that the prevalence of congenital malformation was smaller in the cryopreserved group than in the fresh embryo group. Sutcliffe (1995) studied perinatal outcomes in 91 children from cryopreserved embryos directly after birth. Compared to 83 normally conceived control children, their risk of congenital malformations was 31.9% versus 21.7% for minor and 3.3% versus 2.4% for major congenital malformations, respectively. Both increases were not statistically significant. Olivennes (1996) studied perinatal outcome and 1 to 9 years follow-up in terms of psychomotor development of a group of 82 children born after cryopreservation. No control group was employed. The singleton preterm birth rate and low birthweight rate were 14.7% and 9.3%, respectively. The congenital malformation rate was 3.4%. No pathology was found in psychomotor development. Wennerholm (1997) compared 270 infants after cryopreservation to equal numbers of controls after fresh embryo transfer and after spontaneous conception. No differences were found in preterm or low birthweight rate, nor in malformation rate. In a later study of the same cohort (Wennerholm 1998) no differences were found between 255 children born from cryopreserved embryos, 255 born after IVF with fresh embryos and 252 children from spontaneous pregancies in development nor in the prevalence of chronic illnesses during the first 18 months. Obviously, the number of studies of the effects of cryopreservation on human embryos is limited, the number of subjects per study is rather small and data on follow-up are scant. So far, no evidence exists from human data that cryopreservation may carry risks for the foetus, but clearly more research is needed. A recent French study of long-term effects of embryo freezing in mice showed differences in morphophysiological and behavioural features between cryopreserved and control mice, some of which only appeared at older age. The relevance of these data for the human situation remains to be determined (D'Souza 1997).

Miscellaneous long-term effects

Growth and physical outcome at two years of age was measured in 314 IVF children and compared to that of 150 controls matched for plurality and gestational age (Saunders 1996). IVF status was not an independent factor for physical outcomes, although poorer outcomes were related to the effects of multiple births.

Recently, sleep apnoea in a group of 50 IVF newborns compared to a control group of 50 infants who were all born at term and were matched for gestational age and birthweight has

been reported (Audiens 1995). Sleep apnoea seems to occur more often in babies born from multiple pregancies and with growth retardation (Beal 1989). No difference was found in incidence of apnoea in the IVF versus the control group, although IVF children had significantly more periodic breathing episodes than control children, possibly indicating a more immature respiratory pattern. IVF singleton babies were more prone to breathing abnormalities than IVF twins, which led the authors to conclude the phenomenon is related to IVF per se and not to the infants being part of a multiple pregnancy. The importance of these results cannot yet be decided. It is the first time this relationship has been reported and further, it is unclear from the description of the data whether this is a primary or a secundary analysis. Also, no biological explanation was offered.

Conclusion

Twenty years after the birth of the first IVF baby, many of the effects of the procedure on the children born from it, have become clear, although a substantial number of questions remain unanswered. One of the main problems with IVF to date remains the high incidence of multiple pregnancies. Further, IVF singleton pregnancies and IVF twin pregnancies compared to control singleton or twin pregnancies appear to be at higher risk of adverse outcomes such as preterm birth and low birthweight. Whether this is an effect of the procedure per se, or is related to maternal factors, or a combination of both, remains to be studied. The risk of congenital malformations does at the moment not seem to be elevated. In the future, adequate numbers of subjects should be studied for specific defects, before can be concluded definitively that IVF does not increase the risk of congenital malformation. More data are urgently needed on the possible effects of embryo freezing on the health of the children. Until then, it remains unclear whether embryo freezing is a safe procedure. Psychomotor development of IVF children does not seem to be disturbed, possibly because the quality of IVF parenting is high. Longer follow-up of IVF children is necessary, however, before long-term adverse effects can be excluded. In the future, data on fertility of IVF children will be needed to determine whether the parental fertility problem has been passed on to the next generation.

It is clear that the IVF procedures carry certain short term risks for the offspring. Until further and more extensive studies are conducted, we will not know whether IVF poses additional long term risks for the children.

Table 1 Prematurity (gestational age < 37 weeks), low birthweight (< 2500 grammes) and SGA (Small for Gestational Age, <10th percentile) in IVF singletons

Author and (year)	Number of IVF children studied	number of controls studied	Gestation [1]		Birthweight [2]		SGA (%)		factors controlled for in the analyis
			IVF Children	controls	IVF children	controls	IVF children	controls	
Australian IVF coll. Group (1988)	700	-	18.5%	-	15.5%	-	-	-	-
Beral (1990)	741	-	-	-	12%	-	-	-	-
MRC (1990)	1267	population	13%	6%	12%	-	-	-	-
Doyle (1992)	648	general population	13%	6%	11%	6%	17%	10%	-
Oliviennes (1993)	162	263 * / 5096 **	11.1%	6.1% * / 4.4% **	11.1%	6.5% * / 3.6% **	11.2%	10.6% * / 5.9% **	-
Rufat (1994)	916	-	12.2%	-	12.3%	-	-	-	
FIVAT (1995)	3822	-	9.4%	-	11.2%	-		-	-
Tan (1992)	494	978	14%	8%	14%	7.0%	18.0%	-	maternal age
Verlaenen (1995)	140	140	38.7 weeks	39.8 weeks	3175 gr	3393 gr	-	-	maternal age, height, weight, parity

Author and (year)	Number of IVF children studied	number of controls studied	Gestation [1]		Birthweight [2]		SGA (%)				factors controlled for in the analyis
			IVF Children	controls	IVF children	controls	IVF children		controls		
Petersen (1995)	70	70	7%	4%	16%	0%	-		-		age, parity
D'Souza (1997)	150	150	38.4 weeks	39.5 weeks	3016 gr	3400 gr	-		-		sex of child, social class, gestational age
Wang (1994)	465	21547	13.8%	9%	-	-	16.3%•	5.6%○	10%•	3%○	age, parity
Gissler (1995)	1015	190697	11.0%	4.5%	6.2%	3.7%	-		-		smoking marital status, education, parity
Sanders (1996)	194	111	38.5 weeks	38.7 weeks	3196 gr	3294 gr	-		-		gestational age
Reubinoff (1997)	260	260	8.8%	3.9%	11.2%	1.6%	12.5%		12.9%		maternal age, parity, ethnic origin, location and date of delivery

1) % prematurity, mean gestational age (weeks)
2) % low birthweight, mean birthweight (grammes)
* stimulated cycles
** natural cycles
• all
○ primiparous

Table 2: Behaviour and psychomotor development of IVF children as compared to controls

Author (year)	number of IVF children	number of controls	Age	test used	outcome	controlling for
Mushin (1986)	33	norm population	12-20 months	Bayley Scales of Infant Development	Development within range	
Morin (1989)	83	93	?	Bayley test: 1 mental development index 2 psychomotor development index	no difference on NDI, slightly better outcome for IVF children on PDI:114 vs. 108	age of infant, multiplicity, sex, race, maternal age, parental education and income
Brandes (1992)	116	116	12-45 months	Bayley Scales / Stanford Binet Scales	no difference between the groups	birthweight, gestational age, birth order, mode of delivery, sex, age of child, age of mother
Raoul Duval (1994)	33	33	3 years	Brunes-Lézine	no difference	parity, socio-economic status, mother's age, number of children
Ron-El (1996)	30	30	> 28 months	General Cognitive Index Test	no difference	maternal age, gestational age, birthweight
Cederblad (1996)	99	norm population	33-85 months	Griffith Scale of Mental Development/Child Behaviour Check List	development within range	-

References

American Society for Reproductive Medicine/Society for Assisted Reproductive Technology Registry. Assisted reproductive technology in the United States and Canada. Fertility and Sterility 1995;64:13-21.

Audiens H, Denayer E, Ahlenajafie N, et al. Sleep Apnoea in babies born after in vitro fertilization. European Journal of Pediatrics 1995;154:454-7.

Australian In vitro fertilization Collaborative Group. High incidence of preterm births and early losses in pregnancy after in vitro fertilization. BMJ (Clinical Research ed.) 1985;291:1160-3.

Australian In vitro fertilization Collaborative Group. In vitro fertilization pregnancies in Autralia and New Zealand, 1979-1985. Medical Journal of Australia 1988;148:432-6.

Beal S. Sudden infant death syndrome in twins. Pediatrics 1989;84:1038-44.

Beral V, Doyle P, Tan SL, Mason BA, Campbell S. Outcome of pregnancies resulting from assisted conception. British Medical Bulletin 1990;46:753-68.

Bernasko J, Lynch L, Lapinski R, Berkowitz RL. Twin Pregnancies Conceived by Assisted Reproductive Techniques, Maternal and Neonatal Outcomes. Obstetrics and Gynecology, 1997;89:368-72.

Biggers JD. In vitro fertilization and embryo transfer in human beings. New England Journal of Medicine;1981:304-6.

Botting BJ, Davies IM, Macfarlane AJ. Recent trends in the incidence of multiple births and associated mortality. Archives of Disease in Childhood 1987;62:941-50.

Brandes JM, Scher A, Izkovics J, et al. Growth and development of children conceived by in vitro fertilization. Pediatrics, 1992;90:424-9.

Cederblad M, Friberg B, Ploman F, et al. Intelligence and behaviour in children born after in vitro fertilization treatment. Human Reproduction, 1996;11:2052-7.

Cohen J, Mayaux MJ, Guihard-Moscaro L. Pregnancy outcomes after in vitro fertilisation. A collaborative study on 2342 pregnancies. Annals of the New York Academy of Science, 1988;541:1-6.

Doyle P, Beral V, Maconochie N. Preterm delivery, low birthweight and small-for-gestational-age in liveborn singleton babies resulting from in vitro fertilization. Human Reproduction, 1992;7:425-8.

Dulioust E, Toyama K, Busnel M, et al. Long-term effects of embryo freezing in mice. Proceedings of the National Academy of Sciences of the United States of America, 1995;92:589-93.

D'Souza SW, Rivlin E, Cadman J, et al. Children conceived by in vitro fertilisation after fresh embryo transfer. Archives of Disease in Childhood, 1997;6:F70-F74.

FIVNAT. Bilan FIVAT 1994. Contraception, Fertilité, Sexualité (Paris), 1995;3;490-3.

Friedler S, Mashiach S, Laufer N. Births in Israel resulting from in vitro fertilisation /embryo transfer, 1982-1989. National Registry of the Israeli Association for Fertility Research. Human Reproduction, 1992;7:1159-63.

Friedler S, Mordel N, Lipitz S, et al. Perinatal outcome of triplet pregnancies following assisted reproduction. Journal of assisted reproduction and Genetics, 1994;11,9:459-62.

Gissler M, Malin Silverio M, Hemminki E. In vitro fertilisation pregnancies and perinatal health in Finland 1991-1993. Human Reproduction 1995;10:1856-61.

Golombok S, Brewaeys A, Cook R, et al. The European study of assisted reproduction families: family functioning and child development. Human Reproduction, 1996;11:2324-31.

Howell RJS, Economides D, Teisner B, Farkas AG, Chard T. Placental proteins 12 and 14 in pre-eclampsia. Acta Obstetricia et Gynecologica Scandinavica, 1989;68:237-40.

Jauniaux E, Englert Y, Vanesse M, Hiden M, Wilkin P. Pathologic Features of Placentas From Singleton Pregnancies Obtained by In Vitro Fertilization and Embryo Transfer. Obstetrics and Gynecology, 1990;76:61-4.

Johnson MR, Abbas A, Norman-Taylor JQ, et al. Circulating placental protein 14 (PP14) in the first trimester of spontaneous and IVF pregnancies. Human Reproduction, 1993;8:323-6.
Johnson MR, Irvine R, Hills F, et al. Superovulation, IGFBP-1 and birthweight. European Journal of Obstetrics & Gynaecology and Reproductive Biology, 1995;59:193-5.
Kiely JL, Susser M. Preterm birth, intrauterine growth retardation, and perinatal mortality. American Journal of Public Health, 1992;82:343-5.
Lancaster PAL. Congenital malformations after in vitro fertilization. Lancet 1987;ii:1392-3.
Lino K, Sjoberg J, Seppala M. Elevated circulating levels of a decidual protein, placental protein 12, in preeclampsia. Obstetrics and Gynecology, 1986;68:58-60.
MacMahon CA, Ungerer JA, Tennant C, Saunders D. Psychosocial adjustment and the quality of the mother-child relationship at four months postpartum after conception by in vitro fertilization. Fertility and Sterility, 1997;68:492-500.
Morin NC, Wirth FH, Johnson DH, et al. Congenital malformations and psychosocial development in children conceived by in vitro fertilization. Journal of Pediatrics, 1989;15:222-7.
MRC Working Party on Children Conceived by In Vitro Fertilization. Births in Great Britain resulting from assisted conception 1978-87. BMJ (Clinical Research ed.) 1990;300:1229-33.
Mushin DN, Barreda-Hanson MC, Spensley JC. In vitro fertilization children: early psychosocial development. Journal of In Vitro Fertilisation and Embryo Transfer, 1986;3(4):247-52.
Olivennes F, Kadhel P, Rufat P, et al. Perinital outcome of twin pregnancies obtained after in vitro fertilization: comparison with twin pregnancies obtained spontaneously or after ovarian stimulation. Fertility and Sterility, 1996;66:105-9.
Olivennes F, Schneider Z, Remy V, et al. Perinatal outcome and follow-up of 82 children aged 1-9 years old conceived from cryopreserved embryos. Human Reproduction, 1996;11:1565-8.
Olivennes F, Kerbrat V, Rufat P, et al. Follow-up of a cohort of 422 children aged 6 to 13 years conceived by in vitro fertilization. Fertility and Sterility, 1997;67:284-9.
Olivennes F, Rufat P, André B, et al. The increased risk of complication observed in singleton pregnancies resulting from in vitro fertilization (IVF) does not seem to be related to the IVF method itself. Human Reproduction 1993;8:1297-1300.
Petersen K, Hornnes PJ, Ellingsen S, et al. Perinatal outcome after in vitro fertilisation. Acta Obstetricia et Gynecologica Scandinavica, 1995;74:129-31.
Raoul-Duval A, Bertrand-Servais M, Lenir-Könirsch H, Frydman R. Psychological follow-up of children born after in vitro fertilization. Human Reproduction, 1994;9:1097-1101.
Reubinoff BE, Samueloff A, Ben-Haim M, Friedler S, Schenker JG, Lewin A. Is the obstetric outcome of in vitro fertilized singleton gestations different from natural ones? A controlled study. Fertility and Sterility, 1997;67:1077-83.
Rizk B, Doyle P, Tan SL, et al. Perinatal outcome and congenital malformations in in vitro fertilization babies from the Bourn - Hallam group. Human Reproduction, 1991;6:1259-64.
Rizk B, Edwards RG, Niccolini U, Meagher S, Lamont RF. Edwards' syndrome after the replacement of cryopreserved thawed embryos. Fertility and Sterility 1991;55:208-10.
Ron-El R, Lahat E, Golan A, et al. Development of children born after ovarian superovulation induced by long-acting gonadotropin-releasing hormone agonist and menotropins, and by in vitro fertilization. Journal of Pediatrics, 1994;125:734-7.
Rufat P, Olivennes F, Mouzon de J, Dehan M, Frydman R. Task force report on the outcome of pregnancies and children conceived by in vitro fertilization. Fertility and Sterility, 1994;61:324-30.
Saunders DM, Mathew M, Lancaster PAL. The Australian Register: current research and future role: a preliminary report. Annals of the New York Academy of Science, 1988;160:209-13.
Saunders K, Spensley J, Munro J, Halasz G. Growth and Physical Outcome of Children Conceived by In Vitro Fertilization. Pediatrics, 1996;97:688-92.
Seppala M. The world collaborative report on in vitro fertilization and embryo replacement: current

stat of the art in January 1984. Annals of the New York Academy of Science, 1985:442-558.
Sutcliffe AG, Souza SWD, Cadman J, et al. Minor congenital anomalies, major congenital malformations and development in children conceived from cryopreserved embryos. Human Reproduction, 1995;10:3332-7.
Tan SL, Doyle P, Campbell S, et al. Obstetric outcome of in vitro fertilization pregnancies compared with normally conceived pregnancies. American Journal of Obstetrics and Gynecology, 1992;167:778-84.
Trounson A, Wood C. In vitro fertilization results, 1979-1982, at Monash University, Queen Victoria, and Epworth Medical Centers. Journal of In Vitro Fertilisation and Embryo Transfer, 1984:1-42.
Verlaenen H, Cammu H, Derde MP, Amy JJ. Singleton Pregnancy After In Vitro Fertilization: Expectations and Outcome. Obstetrics and Gynecology, 1995;86:906-10.
Wada L, Macnamee MC, Wick K, Bradfield JM Brinsden PR. Birth characteristics and perinatal outcome of babies conceived from cryopreserved embryos. Human Reproduction, 1994;9:543-6.
Wang JX, Clark AM, Kirby CA, et al. The obstetric outcome of singleton pregnancies following in vitro fertilization/gamete intra-Fallopian transfer. Human Reproduction, 1994;9:141-6.
Wang HS, Lim J, English J, Irvine L, Chard T. The concentration of insulin-like growth factor (IGF-1) and insulin-like growth factor-binding protein-1 (IGFBP-1) in human umbilical cord serum at delivery: relation to fetal weight. Journal of Endocrinology, 1991;129:459-64.
Walker SK, Heard TM, Seamark RF. In vitro culture of sheep embryos without co-culture: successes and perspectives. Theriogenology, 1992;37:111-26.
Wennerholm UB, Albertsson-Wikland K, Bergh C, et al. Postnatal growth and health in children born after cryopreservation as embryos. The Lancet, 1998;351:1085-7.
Wennerholm UB, Hamberger L, Nilsson L, et al. Obstetric and perinatal outcome of children conceived from cryopreserved embryos. Human Reproduction, 1997;12:1819-25.
Willadsen WM, Janzen RE, McAlister RJ, et al. The viability of late morulae and blastocysts produced by nuclear transplantation in cattle. Theriogenology, 1991;35:161-70.

Chapter 3

Preterm birth in pregnancies following in vitro fertilization. Results of a Dutch multicentre prospective study

S.E. Buitendijk[1], J. Koudstaal[2], S. Anthony[1], R.A. Leerentveld[3], H.V. Hogerzeil[4], C.A.M. Jansen[5], R. Brand[1], P.H. Verkerk[1]

1 TNO Prevention and Health, Division of Child Health, PO Box 2215, 2301 CE Leiden, the Netherlands
2 Department of Obstetrics and Gynaecology, University Medical Centre, PO Box 85500, 3508 GA Utrecht, the Netherlands
3 Department of Obstetrics and Gynaecology, Isala klinieken, location Sophia, PO Box 10400, 8000 GK Zwolle, the Netherlands
4 Department of Obstetrics and Gynaecology, Academic Medical Centrum, Meibergdreef 9, 1105 AZ Amsterdam, the Netherlands
5 Department of Obstetrics and Gynaecology, Diaconnessenhuis, PO Box 998, 2270 AZ Voorburg, the Netherlands

Submitted for publication

Abstract

In 2043 IVF pregnancies from which 2636 children were born, the preterm birth rate was studied. It was compared to the preterm birth rate in two separate control groups, the SMOCC- and the LVR-control group. IVF mothers were of lower parity, of higher education and less likely to belong to an ethnic minority than women in the two control groups. They were less likely to have pre-pregnancy diabetes or hypertension. They smoked less and consumed less alcohol than women in the SMOCC-cohort. In the IVF group the rates of induced labour and elective caesarean section were higher than those in the LVR-control group. Women delivered less often at home. Of the IVF pregnancies 26.5% were multiple, compared to 1.6% in both control cohorts. Preterm delivery occurred more often in the IVF cohort, 23.8% vs. 6% in the SMOCC and 7.6% in the LVR-cohort, respectively. Singleton IVF pregnancies ended preterm twice as frequently as control pregnancies did (9.3% vs. 5.2% and 6.9% in the SMOCC and LVR-cohorts, respectively). Singleton IVF children had an increased risk of perinatal death (3.0% vs. 1.5% in the LVR-cohort). Logistic regression analyses were carried out to correct for a large number of variables simultaneously. After correcting for multiplicity, age, parity, ethnicity, mother's education, smoking or alcohol use and pre-pregnancy diabetes and hypertension, in IVF exposed women the odds ratio of the pregnancy ending before 37 weeks was 1.6 (95% CI [1.2 – 2.1]) compared to SMOCC-women. After correcting for multiplicity, age, parity, ethnicity, pre-pregnancy diabetes and hypertension, and whether or not labour was induced, in IVF exposed women the odds ratio of the pregnancy ending before 37 weeks was 1.2 (95% CI [1.1 – 1.4]) compared to LVR-women. In terms of number of days, IVF pregnancies were 2.4 days shorter than SMOCC-pregnancies and 2.3 days shorter than LVR-pregnancies, after correcting for maternal and pregnancy variables in the model. Women suffering from unexplained infertility (which may inherently be related to maternal pathology and chronic conditions negatively influencing pregnancy outcomes) were slightly less likely to suffer preterm birth than women in other indication groups. Idiopathic infertility does not seem to explain the elevated risk.

Introduction

The introduction of In Vitro Fertilisation (IVF) during the early 1980's has not been coupled with formal evaluation of the possible adverse effects on the health of the women, the course of the pregnancies or the health of the babies and children born from it. IVF has nevertheless become widespread. In the Netherlands, the first IVF baby was born in 1983. In 1997, approximately 1 to 1.5 % of all Dutch babies were born after IVF.

Each of the different phases of the IVF treatment could carry a risk of side-effects for the developing embryo (Biggers 1981). First, the hormonal treatment may have potential teratogenic effects. Second, the phase of oocyte retrieval may cause mechanical damage of the egg cells. Third, substances in the culture medium in which eggs and sperm are brought together, may have teratogenic effects on the growing embryo. Fourth, the artificial transfer of the embryo into the uterus may result in a less optimal place of nidation and placentation than may be the case in the natural situation. Fifth, surplus embryos may be frozen to be transferred in a later cycle. Freezing may potentially be harmful to the embryo. It is as yet

unclear whether these potentially damaging effects do lead to pregnancy problems such as miscarriages, preterm births, or growth retardation, congenital malformations or long-term (developmental) problems in IVF children.

IVF pregnancies are more likely to end prematurely than spontaneously conceived pregnancies. In most IVF cycles, more than one embryo is being transferred into the uterus. Consequently, a large percentage (25% to 30%) of IVF pregnancies are multiple pregnancies (American Society for Reproductive Medicine 1995, FIVNAT 1995, D'Souza 1997). Thus, almost half of all children born after IVF are part of a multiple pregnancy. Multiple pregnancies inherently carry a higher risk of complications of which preterm birth, intrauterine growth retardation and perinatal mortality are the most important (Botting 1987, Kiely 1992).

Not only IVF multiple pregnancies, but singleton IVF pregnancies also, appear to be more likely to end preterm (Australian In-Vitro Fertilization Collaborative Group 1985, Saunders 1988, AIHW 1990, Beral.1990, Fiedler 1990, Wennerholm 1991, Doyle 1992, Tan 1992, McFaul 1992, Olivennes 1993, Wang 1994, FIVNAT 1995, Gissler 1995, Tanbo 1995, Verlaenen 1995, Bergh 1999). Whether this phenomenon is related to effects of the IVF procedure per se, to iatrogenic effects or to the IVF mothers' profiles is still a matter of debate. IVF mothers, on average, are older and of lower parity than mothers in the general population. They may differ in other characteristics that are known to influence risk of preterm birth, such as socio-economic status, ethnicity, smoking and drinking behaviour and period of infertility preceding the pregnancy. A number of studies into outcomes of IVF pregnancies use population figures as controls (Australian In-Vitro-Fertilisation Collaborative Group 1988, Doyle 1992, Rufat 1994) or employ no control groups at all (Beral 1990, FIVNAT 1995), thus rendering it impossible to study the effects of maternal characteristics and iatrogenic factors such as preterm induction of labour. In a number of studies into outcomes of IVF pregnancies, attempts have been made to control for a limited number of factors, such as age and parity (Tan 1992, Verlaenen 1995), age, parity and fertility problems (Petersen 1995), age, parity, ethnic origin and place of birth (Reubinoff 1997), smoking, marital status and education (Gissler 1995) or infant sex and social class (D'Souza 1997). These studies showed a 1.5 to 2 times higher preterm rate in the IVF population than in the control group. The numbers in these studies are small. They range from 70 to 800 IVF singletons. A large retrospective cohort study was recently published (Bergh 1999). In this study close to 6000 IVF pregnancies were compared with a large number of population controls (Bergh 1999). After correcting for maternal age and parity and duration of infertility the risk of delivery before 32 weeks and the risk of delivery before 37 weeks showed to be 1.5 times higher in IVF pregnancies than in controls. Although smoking habit and level of education differed between IVF- and controls in this study, they were not included in the analysis studying risk of preterm birth. Caesarean section rates were higher in the IVF than in the control group, but were not taken into account in the preterm birth analysis, thus rendering it impossible to determine what the possible contribution may be of iatrogenic preterm birth on the total preterm birth rate.

From the studies carried out to date it cannot be concluded what the origin is of the increased risk in singleton IVF pregancies. From existing studies it is clear it is partly related to age and parity that differ between IVF- and control mothers. An increased preterm birth rate may also be due to the fact that IVF pregnancies are more prone be induced or delivered by primary caesarean section and therefore end prematurely for iatrogenic reasons. It may further be related to duration and/or type of infertility. A number of studies show that women who have tried longer than 1.5 years to become pregnant (such as IVF mothers), are more likely to have adverse pregnancy outcomes such as miscarriages and preterm births (Ghazi 1991, Williams 1991). In the study by Bergh et al (1999) relative risk of preterm birth indeed decreased after correcting for duration of infertility. This phenomenon may reflect underlying maternal pathology related both to the infertility and to the outcome of the pregnancy, once achieved. Other factors, such as decreased smoking in IVF mothers, or differences in education or ethnicity between IVF- and controls may influence effect estimates and need to be taken into account in the analyses. Last, it may be the IVF procedure itself that increases the risk, through a yet unknown mechanism.

This study is the first, to our knowledge, in which the number of IVF mothers is large enough and information is available on enough factors in the IVF- as well as in the control population, to enable controlling for a substantial number of confounders simultaneously. We were, thus, able to better determine the extent of an increased risk in the IVF group if one existed and to distinguish between procedure-related effects and other (maternal and iatrogenic) effects on risk of prematurity.

Data collection

The study was carried out by TNO Prevention and Health in collaboration with the IVF Working Group of the Dutch Society of Obstetricians and Gynaecologists. The Netherlands has 12 IVF centres holding laboratory facilities, of which 8 are academic and 4 non-academic. In these IVF centres, the entire IVF procedure can be performed, from the (patient) intake to the embryo transfer and the follow-up. The IVF centres collaborate with so-called transport hospitals. In the transport hospitals, all parts of the IVF procedure are being performed with the exception of the laboratory phase and embryo transfer. The laboratory phase is being carried out in one of the IVF centres. Some transport hospitals co-operate with other hospitals called satellite hospitals (Roest et al., 1995). In this construction, three hospitals participate in the treatment procedure. The harvesting of the oocytes takes place in the transport hospital, the laboratory phase and embryo transfer are being performed in the IVF centre and the other parts of the procedure i.e., patient selection, hyperstimulation, cycle monitoring and follow-up after embryo transfer are being carried out in the satellite hospital. In the Dutch IVF program, 12 IVF centres, 23 transport hospitals and 6 satellite hospitals take part. All hospitals except one transport hospital agreed to participate in our study.

From January 1st 1994, to December 31st 1995, women who were interested in participating were asked written informed consent. It was obtained shortly after the IVF treatment had shown to have resulted in an ongoing pregnancy, diagnosed primarily with fetal heart action

at ultrasound. Women who subsequently miscarried before 16 weeks of pregnancy, were excluded from the study. Data on infertility history and treatment schedule were obtained through the participating IVF clinic. Two months after the expected date of delivery the IVF patient received a mailed questionnaire containing questions about her obstetric and medical history as well as the course and outcome of her pregnancy. The provider of the prenatal care (general practitioner, midwife or obstetrician) was asked to provide information concerning the prenatal period and delivery, through a questionnaire. If the newborn had been in paediatric care a copy of the paediatrician's report was obtained.

Control groups

Two control groups were employed. First, pregnancies from the Social Medical Survey of Children attending Child Health Clinics (SMOCC) served as controls (Herngreen1992). In the SMOCC-study, data were collected through postpartum interview with all mothers from liveborn infants in the areas of 21 different child health clinics. Of 2151 babies born between April 1st 1988 and October 31st 1989, data were gathered concerning maternal, socio-demographic and health characteristics, the course of the pregnancy and the health of the newborn.
Since the SMOCC-dataset is restricted to live-born babies only, a second control dataset was employed that also contains information on infants that died in the perinatal period. In order to create this second control group representative of all births in the Netherlands, the 1995 data of two Dutch national databases were combined: the National Perinatal Database of obstetrical care delivered by midwives and general practitioners (LVR-1) and the National Perinatal Database of obstetrical care delivered by obstetricians (LVR-2). Midwives, GP's and obstetricians register information about pregnancy, delivery and puerperium of pregnancies with a gestational age of at least 16 completed weeks. In 1995, 89% of all midwives and obstetricians participated in the LVR. The GP's participation was very limited. Records in the joint database do not have a unique number. Further, pregnancies (and children) can occur twice in the joint database in case women are referred during pregnancy from a midwife to an obstetrician. An aggregated perinatal data file representative of all 1995 births in the Netherlands was created. In order to identify each anonymous child as a single record, for each record a unique key was defined. A number of identifying variables were used such as mother's postal code and child's date of birth. This dataset has been used in a study on perinatal deaths in the Netherlands and has been validated by comparing the number of deaths in this database to the number in the 1995 perinatal database of the Dutch Central Bureau of Statistics (CBS). Methods are described in detail elsewhere (Anthony, in preparation). As a last step, an attempt was made to identify IVF pregnancies in the LVR-dataset. A number of key variables among which mother's birth date, birth date of the child, sex of the child were used to identify the IVF pregnancies in the LVR-dataset. Also, a variable in the LVR-dataset indicating whether IVF treatment had taken place, was used to identify IVF pregnancies. Those pregnancies that likely were IVF pregnancies, were subsequently removed from the LVR-control dataset.

Non-participation

During the data collection period, not all IVF clinics that originally agreed to participate were able to obtain consent from all eligible women in their center. Of the total of 41 centers, 5 did not contribute any patient data to the study, or less than 5% of their potential. Of those, 4 were transport centers and 1 was a satellite center. Of the remaining 36, 11 centers (1 center with laboratory facilities and 10 transport hospitals) obtained consent from less than 80% of their eligible patients. 25 Centers obtained informed consent from between 80% and 100% of their pregnant IVF patients. During the course of the study, further inquiry was made by the research team into reasons for non-participation. This inquiry revealed that non-participation was in all centers related to either logistic problems or staff problems such as pregnancy or sick leave or depature of the physician coordinating the data collection. Judging from the reasons for the non-participation, it seems unlikely this introduced a bias. First, IVF centers in the Netherlands draw patients based mostly on region of residence and not based on risk profile, age or cause of infertility. It is therefore unlikely that non-participating centers had patient populations with worse or better pregnancy outcomes than participating centers. Second, in centers that participated partially, logistic problems instead of patient refusal appeared to have led to incomplete data collection. It is therefore unlikely that non-participating patients had a different risk profile and pregnancy outcome than did participating patients. Overall, patients proved very willing to participate. The refusals that did occur were often related to language problems.

Definition of main (outcome) variables

Duration of pregnancy is defined differently in the IVF- than in the control cohorts. In the SMOCC and LVR-cohort, pregnancy duration is based on the mother's information on first day of last menstrual period, or in absence of that information, on data from ultrasound examination. Gestational age at birth is defined as the number of days between first day of last menstrual period and date of birth. In the IVF cohort, pregnancy duration is based on day of embryo transfer. Date of embryo transfer was subtracted from the birth date of the baby. In order to approximate number of days of pregnancy duration as defined in the control cohorts, 17 days were added to the result of this subtraction. Thus, in the IVF cohort, gestational age at birth is defined as the period between date of embryo transfer and delivery, plus 17 days. The potential problem with this approach is that there can be no certainty as to the exact day in a woman's natural cycle the embryo descends into the uterus to nidate (Goldenberg 1997). In the calculation for our IVF cohort, it is being assumed that in a natural cycle the fertilised egg travels to the uterus in 3 days and that ovulation, on average, takes place 14 days after the first day of the last menstrual period. The 3 days are more likely an overestimation than an underestimation. Preterm birth was defined as birth before 37 completed weeks. Very preterm birth was defined as birth before 32 completed weeks. Small-for gestational age (SGA) was defined as birthweight below the tenth percentile of the Dutch reference curve (Kloosterman 1970). Low birthweight (LBW) was defined as birthweight < 2500 g. An elective cesarean section was defined as an cesarean section that is performed before the onset of labor. Since adding 17 days may result in an overestimation of a negative effect of IVF, we also defined IVF gestational age by adding 15 and 16 days, respectively.

Data on maternal and pregnancy variables were collected in a number of ways. In the IVF group, data were collected through the LVR-record form, the physician's medical correspondence and a questionnaire mailed to the mother during the first weeks after birth. To compare data in the IVF cohort to those in the LVR-controlgroup, demographic and pregnancy variables for the IVF group were defined and collected similarly to those in the LVR-group, namely through the IVF form or the standard medical correspondence which often follows the LVR-format. Only in case of missing IVF data, information from the mother's questionnaire was used. For comparisons between the IVF and SMOCC groups, data on smoking, alcohol use, medication use, educational level, ethnicity and maternal height as well as those on condition of the infant, were collected for the IVF women in a similar fashion as in the SMOCC- data collection, namely primarily through the mother's questionnaire. In the SMOCC study data were collected after birth, although in this case by in-person interview. In both cases, however, data collection through the mother's questionnaire was retrospective and around the same period after birth. Variable definition was the same in the IVF- and the SMOCC-cohorts.

Analyses

The relationship between IVF exposure on the one hand and maternal demographic and lifestyle variables, pregnancy and delivery characteristics and neonatal outcome variables on the other hand, were tested in bivariate analyses, using Chi-square analyses for categorical and Student's t-test for continuous variables. Main outcome variables such as birthweight were checked for differences in patterns of digit preference between the three groups. Logistic regression analyses were carried out in order to differentiate between the effect of maternal demographic variables (such as age and parity), maternal lifestyle variables, pre-existent diabetes and hypertension and the effect of IVF exposure per se, on the risk of preterm birth. Variables were added to the model in a stepwise fashion. Similarly, linear regression analyses were carried out to test for the effect of these variables on pregnancy duration in days. Possible interaction effects between IVF exposure and maternal age as well as smoking, were also included in the models. A number of key analyses were re-run with IVF pregnancy duration defined as birth date of the baby minus date of embryo transfer, plus 15 and plus 16 days instead of plus 17 days.
Analyses were carried out comparing the risk of preterm birth for the four different subgroups of indications for IVF treatment (tubal pathology, sperm abnormalites, idiopathic infertility and combinations).
Last, analyses were carried out comparing IVF- and control pregnancies after restricting both the IVF- and the control groups to low-risk subgroups homogeneous on all factors but the IVF exposure.
In the comparison between IVF and SMOCC a subgroup was constructed consisting of singleton pregnancies in women pregnant for the second time, who were between the ages of 20 and 35, were not belonging to an ethnic minority, were non-smoking, non-drinking, and had no pre-existent diabetes or hypertension. In the comparison between IVF and LVR a subgroup was contructed consisting of singleton pregnancies in women pregnant for the second time, who were between the ages of 20 and 35, were not belonging to an ethnic

minority and had no pre-existent diabetes or hypertension. Thus, if covariates were measured imperfectly, they are less likely at this low level of risk and in these 'pure' subsamples, to alter the effect of IVF exposure on the outcomes measured and result in a spurious association between IVF exposure and preterm birth.

In order to ensure that non-participation did not bias the results of our study, it was decided that data should be analysed in two ways. First, the entire dataset, including patient data from 'incomplete' centers, should be analysed. Second, the analysis was to be repeated with data only from the centers that provided information on over 80% of their pregnant IVF patients. If similar outcomes were to occur in both groups, it may be assumed the non-participation did not bias the results of our study.

Results

During 1994 and 1995, informed consent forms were signed by 2043 women with an IVF pregnancy of at least 16 weeks duration. Of 1979 of these women data on the course of pregnancy and delivery and the condition of the baby were obtained through a written questionnaire filled out by the woman. For 1877 women, similar data were obtained from the midwife or obstetrician through written questionnaires and copy sheets of data from the national obstetric data base. Data on the treatment cycle preceding the IVF pregnancy were available for 1914 of the women in our database. Of the 2043 pregnancies, 2636 children were born, 1502 of whom were singletons and 1134 part of a multiple pregnancy (table 1).

Table 1 Number of subjects in IVF, SMOCC and LVR#-cohorts*

	mothers	all children	children, part of multiple pregnancy	children, singletons
IVF	2043	2636	1134	1502
SMOCC	2119	2151°	65	2061
LVR	156.473	158.992	5002	153.990

* (liveborn) infants from the Social Medical Survey on Children Attending Child Health Clinics

\# National Perinatal database

° of 25 children it could not be determined whether they are part of a multiple or a singleton pregnancy

Table 2 shows the distribution of maternal variables in the IVF cohort as compared to the two control cohorts (SMOCC and LVR). Of IVF women 40.9% were 35 years or older, compared to 10.4% and 13.8% in the SMOCC- and LVR-control groups, respectively. Further, IVF mothers were of lower parity, of higher education and less likely to belong to an ethnic minority than women in the two control cohorts. IVF women smoked less and consumed less alcohol than women in the SMOCC cohort. They were less likely to have pre-pregnancy hypertension and pre-pregnancy diabetes compared to women in the SMOCC-cohort.

Table 2 Characteristics of women in the IVF- and the two control cohorts

	IVF (n = 2043) % (n)	SMOCC# (n = 2119) % (n)	LVR II (n = 156473) % (n)	p-value IVF vs SMOCC	p-value IVF vs LVR
Maternal age (yrs)					
<25	0.7 (14)	16.1 (337)	13.3 (20727)	< 0.0001	< 0.0001
25-29	13.7 (278)	40.5 (846)	35.4 (55260)		
30-34	44.7 (907)	32.9 (687)	37.5 (58541)		
35-39	35.1 (713)	9.3 (195)	12.2 (19075)		
≥ 40	5.8 (118)	1.1 (24)	1.6 (2471)		
Average in height (cm) ± SD	169.0 ± 6.7	168.3 ± 6.9	- *	0.001	-
Ethnicity					
Dutch	93.3 (1844)	91.1 (1877)	83.2 (130044)	0.009	< 0.0001
non-Dutch	6.7 (133)	8.9 (184)	16.8 (26312)		
Education					
below university degree	73.4 (1404)	83.5 (1687)	- *	< 0.0001	-
university degree	26.6 (509)	16.5 (333)	- *		
Parity					
primiparous	69.3 (1415)	42.5 (876)	44.2 (68510)	< 0.0001	< 0.0001
multiparous	30.7 (627)	57.5 (1184)	55.8 (86359)		
Previous miscarriages					
yes	26.0 (531)	18.0 (370)	- *	< 0.0001	-
no	74.0 (1511)	82.0 (1690)	- *		
Pre-existent diabetes					
yes	0.3 (7)	1.1 (22)	0.2 (287)	0.006	0.09
no	99.7 (2015)	98.9 (2038)	99.8 (156141)		
Pre-existent hypertension					
yes	1.6 (32)	2.2 (45)	0.5 (836)	0.16	< 0.0001
no	98.4 (1989)	97.8 (2015)	99.5 (155592)		
Smoking during pregnancy					
no smoking	77.1 (1509)	74.1 (1524)	- *	0.02	-
≥ 1 cigarette/day	22.9 (447)	25.9 (534)	- *		
Alcohol use during pregnancy					-
no alcohol use	84.6 (1651)	74.1 (1525)	- *	< 0.0001	
≥ 1 glass/week	15.4 (301)	25.9 (533)	- *		

* no data available

\# liveborn infants from the Social Medical Survey on Children Attending Child Health Clinics

II National Perinatal database

Totals may vary because of missing values

Table 3 shows the course of pregnancy and delivery in IVF women as compared to women in the control cohorts. In the IVF group the rates of induced labor and of elective cesarean section were higher than those in the LVR control group (two and three times, respectively). In the IVF group women delivered less often at home (5.4% compared to 27.4% in the LVR-control group). Of the IVF pregnancies 26.5% were multiple, compared to 1.6% in both control cohorts.
Preterm delivery occurred more often in the IVF cohort, 23.8% vs. 6.0% in the SMOCC- and 7.6% in the LVR-cohort, respectively. IVF babies were more likely to be growth retarded and to die in the perinatal period (table 4).

Table 3 Course of pregnancy and delivery in women in the IVF- and control cohorts (singleton and multiple pregnancies combined)

	IVF (n = 2043) % (n)	SMOCC (n = 2119) % (n)	LVR (n = 156473) % (n)	p-value IVF vs SMOCC	p-value IVF vs LVR
Onset of labour					
spontaneous	59.9 (1037)		80.4 (125688)	-	< 0.0001
induced	27.5 (476)	-*	15.3 (23866)		
planned c. section	12.6 (218)		4.3 (6690)		
Place of birth					
home	5.4 (106)	-*	27.4 (42845)	-	< 0.0001
hospital	94.6 (1870)		72.6 (113551)		
Birth attendant					
midwife / GP	15.0 (296)	-*	39.1 (61204)	-	< 0.0001
obstetrician	85.0 (1676)		60.9 (95269)		
Number of children					
singleton birth	73.5 (1502)	98.4 (2061)	98.4 (153990)		
twins	24.0 (489)	1.6 (33)	1.6 (2438)	< 0.0001	< 0.0001
triplets	2.5 (52)	0 (0)	0 (43)		
quadruplets	0 (0)	0 (0)	0 (2)		
Duration of pregnancy (weeks)					
< 29	3.9 (78)	0.3 (7)	1.1 (1730)	< 0.0001	< 0.0001
29 – 36	19.9 (403)	5.7 (120)	6.5 (10152)		
≥ 37	76.2 (1540)	93.9 (1966)	92.4 (144115)		
Mean duration of pregnancy (days)	266	278	276	< 0.0001	< 0.0001

* no data available
Totals may vary because of missing values

Table 4 Condition of children in the IVF- and control cohorts (singleton and multiple pregnancies combined)

	IVF (n = 2517/2636)** % (n)	SMOCC (n = 2151) % (n)	LVR (n = 158992) % (n)	p-value IVF vs SMOCC	p-value IVF vs LVR
Sex					
- boys	51.9 (1301)	49.3 (1048)	51.4 (81381)	0.07	0.46
- girls	48.1 (1204)	50.7 (1078)	48.6 (76952)		
Birthweight (grammes)					
< 2000	13.4 (335)	1.8 (39)	3.5 (5556)	< 0.0001	< 0.0001
2000 – 2499	17.7 (441)	4.1 (88)	4.2 (6727)		
2500 – 2999	22.5 (562)	13.7 (292)	14.8 (23561)		
3000 – 3499	25.0 (624)	31.8 (676)	34.3 (54512)		
3500 – 3999	15.7 (392)	34.7 (738)	30.4 (48344)		
≥ 4000	5.6 (139)	13.7 (292)	12.7 (20187)		
Growth retardation					
< p10	19.3 (478)	10.2 (213)	10.6 (16384)	< 0.0001	< 0.0001
- yes	80.7 (1996)	89.8 (1879)	89.4 (138807)		
- no					
Growth retardation < p2.3					
- yes	4.5 (111)	1.8 (37)	2.2 (3438)	< 0.0001	< 0.0001
- no	95.5 (2363)	98.2 (2055)	97.8 (151753)		
Congenital malformation (according to midwife/ physician, directly after birth)					
- yes	4.0 (93)	3.6 (76)*	2.5 (4025)	0.5	< 0.0001
- no	96.0 (2232)	96.4 (2016)*	97.5 (154967)		
Mortality					
- no	94.8 (2205)	-*	98.3 (156351)	-	< 0.0001
- yes, before/during birth	3.0 (70)		1.2 (1985)		
- yes, after birth	2.2 (52)		0.4 (656)		

* no data available

** for comparisons with SMOCC-cohort only liveborn (n = 2517) IVF children were included, for comparison with LVR live- and stillborn (n = 2636) children were considered

Totals may vary because of missing values

In singleton IVF pregnancies also, labour was more often induced and birth more often took place in the hospital, than in singleton control pregnancies. Singleton IVF pregnancies twice as frequently ended prematurely than singleton control pregnancies (9.3% compared to 5.2% and 6.9% in the SMOCC and LVR-cohort, respectively) (table 5). Singleton IVF children had an increased risk of dying in the perinatal period (table 6).

Table 5 Course of pregnancy and delivery in IVF- and control cohorts, singletons only

	IVF (n=1465/1502)** % (n)	SMOCC (n = 2061) % (n)	LVR (n = 153990) % (n)	p-value	
				IVF vs SMOCC	IVF vs LVR
Onset of labour					
- spontaneous	64.8 (803)		80.8 (124252)		< 0.0001
- induced	25.3 (314)	-*	15.0 (23136)	-	
- planned c.section	9.8 (122)		4.1 (6377)		
Place of birth					
- home	7.3 (106)	-	27.8 (42838)	-	< 0.0001
- hospital	92.7 (1350)		72.2 (111075)		
Birth attendant					
- midwife/GP	19.4 (282)	-	39.7 (61193)	-	< 0.0001
- obstetrician	80.6 (1171)		60.3 (92797)		
Duration of pregnancy (weeks)					
< 29	0.4 (6)	0.2 (5)	1.0 (1579)	< 0.0001	< 0.0001
29 – 36	8.9 (129)	5.0 (103)	5.9 (9057)		
≥ 37	90.7 (1314)	94.8 (1952)	93.1 (142883)		

* no data available

** for comparisons with SMOCC-cohort only liveborn (n = 1465) IVF children were included, for comparison with LVR live- and stillborn (n = 1502) children were considered

Totals may vary because of missing values

Digit preference was assessed for birthweight. It occurred only for 5 and 10 grams groups and in a similar fashion in IVF-, SMOCC- and LVR-groups.

Next, a logistic regression model was used to differentiate between the effect of IVF per se and the effect of factors known to be different between IVF- and control groups, such as multiplicity of the pregnancy, maternal age, parity and maternal smoking, on the course of pregnancy and on the condition of the newborn.

*Table 6 Condition of children in the IVF- and control cohorts, **singletons only***

	IVF (n = 1465/1502)** % (n)	SMOCC (n = 2061) % (n)	LVR (n =153990) % (n)	p-value IVF vs SMOCC	p-value IVF vs LVR
Sex					
- boys	52.3 (761)	49.2 (1014)	51.4 (78854)	0.07	0.29
- girls	47.7 (693)	50.8 (1047)	48.6 (74500)		
Birthweight (grammes)					
< 2000	4.1 (60)	1.3 (27)	2.8 (4331)	< 0.0001	< 0.0001
2000 – 2499	5.9 (86)	3.3 (69)	3.5 (5329)		
2500 – 2999	19.2 (278)	13.1 (269)	14.3 (21963)		
3000 – 3499	35.2 (509)	32.4 (667)	35.0 (53839)		
3500 – 3999	26.1 (378)	35.7 (736)	31.4 (48248)		
≥ 4000	9.5 (137)	14.2 (292)	13.1 (20178)		
Growth retardation (<P10)					
- yes	13.8 (199)	9.7 (196)	10.0 (14996)	< 0.0001	< 0.0001
- no	86.2 (1244)	90.3 (1831)	90.0 (135488)		
Congenital malformation (according to midwife or physician)					
- yes	4.3 (57)	3.6 (72)	2.5 (3819)	0.30	< 0.0001
- no	95.7 (1284)	96.4 (1955)	97.5 (150171)		
Mortality					
- no	97.0 (1270)	-*	98.5 (151645)	-	< 0.0001
- yes, before/during birth	2.1 (27)		1.2 (1794)		
- yes, after birth	0.9 (12)		0.4 (551)		

* no data available

** for comparisons with SMOCC-cohort only liveborn (n = 1465) IVF children were included, for comparison with LVR live- and stillborn (n = 1502) children were considered

Totals may differ because of missing values

The tables show a number of different models. First, in table 7 in a full model the effect of IVF exposure as well as the separate effects of all other variables is shown for the SMOCC dataset as a control group, with risk of birth before 37 weeks as outcome. It can be seen that the largest single risk factor is multiple pregnancy. After controlling for all covariates, including multiplicity, the OR for IVF exposure is 1.6 (95%CI [1.2 – 2.1]). Tables 8 through 11 show models in which in a stepwise fashion an increasing number of variables have been taken into account. The outcomes studied are birth before 37 completed weeks of pregnancy, birth before 32 completed weeks of pregnancy and duration of pregnancy in days. It is shown in table 9 that when a woman is IVF exposed, the odds of her pregnancy ending before 37 weeks are 5.5 (95% CI [5.0-6.0]) compared to women in the LVR-cohort.

Table 7 Logistic regression analysis of the effect of IVF exposure and other variables on the risk of birth < 37 weeks (SMOCC – control group)

Variable	n (%)	Odds Ratio [95% CI]	p-value
IVF exposure			
yes	2328 (53.3)	1.6 [1.2 – 2.1]	0.002
no	2040 (46.7)	1.0	
Multiple pregnancy			
yes	1045 (23.9)	14.4 [11.6 – 17.8]	< 0.0001
no	3323 (76.1)	1.0	
Maternal age (years)		1.0 [0.98 – 1.03]	0.95
Maternal height (in cm)		0.98 [0.96 – 0.99]	0.001
Parity			
primiparous	2490 (57.0)	1.1 [0.9 – 1.4]	0.3
multiparous	1878 (43.0)	1.0	
Pre-existent diabetes			
yes	28 (0.6)	5.0 [2.0 – 12.3]	0.0005
no	4340 (99.4)	1.0	
Pre-existent hypertension			
yes	80 (1.8)	2.0 [1.0 – 3.7]	0.04
no	4288 (98.2)	1.0	
Ethnicity			
Dutch	4060 (92.9)	1.0	0.2
non-Dutch	308 (7.1)	0.8 [0.5 – 1.1]	
Maternal education			
non-university	3436 (78.7)	1.2 [1.0 – 1.6]	0.08
university	932 (21.3)	1.0	
Smoking			
≥ 1 cig / day	1042 (23.9)	1.0 [0.8 – 1.3]	0.8
No smoking	3326 (76.1)	1.0	
Alcohol use			
≥ 1 glass / week	877 (20.1)	0.7 [0.5 – 0.9]	0.01
no alcohol use	3491 (79.9)	1.0	

After correcting for multiplicity the odds ratio is 1.5 (95% CI [1.4-1.7]), after also taking into account maternal age, parity and ethnicity the odds ratio is 1.3 (95% CI [1.2 - 1.5]) and after adding pre-pregnancy diabetes and hypertension and whether or not labour was induced or a primary caesarean section took place, the odds ratio is 1.2 (95% CI [1.1 - 1.4]). It is shown in table 8 and 9 that after correction for all variables in the model, the odds ratio for preterm birth < 32 weeks is 2.0 (95% CI [0.9 – 4.1]) for IVF women compared to SMOCC-controls and OR = 0.9 (95% CI [0.7 – 1.1]) compared to LVR-controls. More maternal variables could be included in the SMOCC-models than in the LVR-models. On the other hand, information on induction of labour or elective caesarean section was not available in the SMOCC database. Table 10 and 11 show similar analyses for gestational age in number of days as outcome. IVF pregnancies are 12 and 14 days shorter in duration than SMOCC- and LVR- pregnancies, respectively. After including all available covariates in the model, the effect of IVF was a shortening of 2.4 days pregnancy duration (95% CI [-3.6 to -1.2 days]) and 2.3 days (95% CI [-3.0 to -1.5 days]) compared to SMOCC and LVR, respectively.

Table 8 Stepwise logistic regression analysis of relationship between IVF exposure and preterm birth (< 37 weeks) IVF vs SMOCC-control group

Variables in model		Odds ratio [95% CI#] birth < 37 wk IVF exposed vs non-exposed	Odds ratio [95% CI#] birth < 32 wk IVF exposed vs non-exposed
IVF exposure	*	5.4 [4.5 – 6.6]	5.8 [3.3 – 10.3]
IVF exposure, multiplicity	**	1.6 [1.2 – 2.0]	1.9 [1.0 – 3.7]
IVF exposure, multiplicity, age, maternal height, parity, ethnicity, mother's education	***	1.6 [1.2 – 2.1]	2.0 [1.0 – 4.3]
IVF exposure, multiplicity, age, maternal height, parity, ethnicity, mother's education, smoking, alcohol use, pre-existent hypertension, pre-existent diabetes	****	1.6 [1.2 – 2.1]	2.0 [1.0 – 4.3]

* : model shows effect of IVF exposure only
** : model shows effect of IVF exposure after correcting for multiple pregnancy
*** : model shows effect of IVF exposure after correcting for multiple pregnancy and maternal demographic variables
**** : model shows effect of IVF exposure after correcting for multiple pregnancy, maternal demographic variables, pregnancy variables and other exposures
# : 95% confidence interval

Table 9 Stepwise logistic regression analysis of relationship between IVF exposure and preterm birth (<37 weeks), LVR-control group

Variables in model		Odds ratio [95% CI#] birth < 37 wk IVF exposed vs non-exposed	Odds ratio [95% CI#] birth < 32 wk IVF exposed vs non-exposed
IVF – exposure	*	5.5 [5.0 – 6.0]	3.6 [3.1 – 4.3]
IVF exposure, Multiplicity	**	1.5 [1.4 – 1.7]	1.1 [0.9 – 1.4]
IVF – exposure, multiplicity, age, parity, ethnicity	***	1.3 [1.2 – 1.5]	1.0 [0.8 – 1.2]
IVF – exposure, multiplicity, age, parity, ethnicity, pre-existent hypertension, pre-existent, diabetes, labour induced	****	1.2 [1.1 – 1.4]	1.0 [0.8 – 1.2]

* : model shows effect of IVF exposure only
** : model shows effect of IVF exposure after correcting for multiple pregnancy
*** : model shows effect of IVF exposure after correcting for multiple pregnancy and maternal demographic variables
**** : model shows effect of IVF exposure after correcting for multiple pregnancy, maternal demographic variables, pre-existent chronic diseases and induction of labour
# : 95% confidence interval

Table 10 Stepwise linear regression analysis of the relationship between IVF exposure and duration of pregnancy in days, IVF vs SMOCC

Variables in model		Difference in number of days (95% CI) pregnancy duration IVF exposed vs non-exposed
IVF exposure	*	-12.0 [-13.1 to –11.0]
IVF exposure, multiplicity	**	-2.7 [-3.7 to -1.7]
IVF exposure, multiplicity, age, maternal height, parity, ethnicity, mother's education	***	-2.6 [-3.8 to -1.4]
IVF exposure, multiplicity, age, maternal height, parity, ethnicity, mother's education, smoking, alcohol use, pre-existent hypertension, pre-existent diabetes	****	-2.4 [-3.6 to -1.2]

* : model shows effect of IVF exposure only
** : model shows effect of IVF exposure after correcting for multiple pregnancy
*** : model shows effect of IVF exposure after correcting for multiple pregnancy and maternal demographic variables
**** : model shows effect of IVF exposure after correcting for multiple pregnancy, maternal demographic variables, pregnancy variables and other exposures
# : 95% confidence interval

Table 11 Stepwise linear regression analysis of the relationship between IVF exposure and duration of pregnancy in days, IVF vs LVR

Variables in model		Difference in number of days (95% CI) pregnancy duration IVF exposed vs non-exposed
IVF – exposure	*	-14.1 [-14.9 to -13.3]
IVF exposure, multiplicity	**	-3.3 [-4.1 to -2.6]
IVF – exposure, multiplicity, age, parity, ethnicity	***	-2.9 [-3.7 to -2.1]
IVF – exposure, multiplicity, age, parity, ethnicity, pre-existent hypertension, pre-existent, diabetes, labour induced	****	-2.3 [-3.0 to -1.5]

* : model shows effect of IVF exposure only
** : model shows effect of IVF exposure after correcting for multiple pregnancy
*** : model shows effect of IVF exposure after correcting for multiple pregnancy and maternal demographic variables
**** : model shows effect of IVF exposure after correcting for multiple pregnancy, maternal demographic variables, pre-existent chronic diseases and induction of labour
# : 95% confidence interval

Further, we employed logistic and linear regression models varying the number of days of gestational age in the IVF group as described under 'analyses' adding 16 days and 15 days, respectively to the period between date of embryo transfer and delivery. These results show that if we use a less conservative estimate of the number of days of gestation in the IVF cohort, the estimates of the effect of IVF on the risk of preterm birth before 37 weeks, go up (table 12).

The IVF- and SMOCC-cohorts were compared on risk of preterm birth and duration of pregnancy in a restricted low-risk subsample of singleton pregnancies in women pregnant for the second time, who were between the ages of 20 and 35, were not belonging to an ethnic minority, were non smoking, non alcohol drinking, had no pre-existent chronic conditions and had not previously miscarried. The subsample contained 260 SMOCC-women and 88 IVF women. The risk estimate for preterm birth before 37 weeks was 5.6 (95% CI [2.0 to 16.0]) and duration of pregnancy was 6.2 days shorter in the IVF group (95% CI [-9.3 to – 3.2]). We corrected for maternal age, height and education in the restricted sample. This yielded an odds ratio for preterm birth of 5.6 (95% CI [1.6 – 19.9]) and a difference in number of days of –5.7 (95% CI [-9.2 to –2.2]). Similarly, the IVF- and LVR-cohorts were compared in a subsample of singleton pregnancies in women pregnant for the second time, aged 20 to 35 years, who were not belonging to an ethnic minority and had not previously suffered diabetes or hypertension. Over 41000 LVR-women and 215 IVF women remained in the subsample. The risk estimate for preterm birth before 37 weeks was 3.0 (95% CI [1.9 - 4.6]) and duration of pregnancy was 6.1 days shorter in the IVF group (95% CI [-8.2 to – 4.2]). We corrected for maternal age and whether or not labour was induced in the restricted sample. This slightly reduced the odds ratio for effect of IVF to 2.6 (95% CI [1.7 – 4.1]) and reduced the difference in number of days to –6.2 (95% CI [-8.4 to –4.1]).

Last, we wanted to test the hypothesis that the effect of IVF exposure on the risk of premature labour may be different for different indications. If indeed period of infertility has an effect on pregnancy outcome through underlying maternal pathology, it may be assumed that IVF women with blocked tubes, who are 'fertile' but for the problem with the pathway, are less likely to deliver prematurely than other IVF women. IVF women with idiopathic infertility should be most at risk for unfavourable pregnancy outcomes. Also, women with subfertile partners, or with a combination of indication categories, should be more likely to be subfertile than women with 'pure' tubal pathology. Therefore, we carried out analyses to show mean birthweight and mean duration of pregnancy for different indications. The first indication group was defined as women with tubal pathology only (i.e. with no secondary fertility problem). The second subgroup was defined as women with partners with sperm pathology, the third as women with idiopathic infertility or unsuccesful donor insemination and the fourth as women with a combination of factors. The results are shown in table 13. Within the IVF group we subsequently employed a logistic regression model calculating the risk of birth before 37 completed weeks as well as pregnancy duration in days, respectively, for the four indication categories, while controlling for other factors. Table 14 shows that the odds of preterm delivery do not differ between the different indication categories, indicating that type of infertility does not influence the risk of preterm delivery. Indication category also does not influence pregnancy duration in days (results not shown).

Analyses were repeated for the subgroup of participating centers that contributed over 80% of their patients to our study. The outcomes changed very marginally, resulting in decimal point changes in risk estimates only. We concluded that non-participation did not introduce a bias and have subsequently used all records (both from "complete" and from "incomplete" centers) in our analyses.

Table 12 Estimates of risk of delivery < 37 weeks, for different calculations of gestational age in the IVF cohort vs SMOCC

Variables		Odds ratio [95% CI#]< 37 wk IVF exposed vs non-exposed
IVF exposure[I]	*	1.6 [1.2 – 2.1]
IVF exposure[II]	*	1.8 [1.4 – 2.4]
IVF exposure[III]	*	2.1 [1.6 – 2.8]

\# : 95% confidence interval

* : model shows effect of IVF exposure after correcting for multiplicity, age, maternal height, parity, ethnicity, mother's education, smoking, alcohol use, pre-existent hypertension, pre-existent diabetes

I : gestational age at birth in the IVF cohort defined as period between date of embryo transfer and delivery, plus 17 days

II : same, plus 16 days

III : same, plus 15 days

Table 13 Adverse pregnancy outcomes by indication for IVF treatment, total group

Outcome	Tubal pathology (n = 662)	Sperm abnorm. (n = 476)	Idiopathic infertility (n = 456)	Combination (n = 806)
Birthweight mean, ± SD	2783 (863)	2810 (873)	2740 (910)	2748 (878)
Duration of pregnancy in days mean, ± SD*	261 (28)	264 (27)	261 (29)	261 (30)

* SD = Standard Deviation

Table 14 Logistic regression analysis comparing the risk of preterm delivery between the different indication groups, total group

	n** (%)	Odds ratio [95% CI]# birth < 37 wk		P-value
Indication *				
tubal pathology	600 (27.1)	1.0		
sperm abnormalities	439 (19.8)	0.7	[0.5 – 1.0]	0.05
idiopathic infertility	422 (19.0)	0.7	[0.5 – 1.0]	0.03
combination	757 (34.1)	0.9	[0.7 – 1.2]	0.6

\# : 95% confidence interval

* : controlling for multiplicity, age, maternal height, parity, ethnicity, mother's education, smoking, alcohol use, pre-existent hypertension, pre-existent diabetes

** : totals may differ because of missing values

Discussion

Our study results are in accordance with those from other recent studies into the risk of preterm birth in IVF pregnancies (American Society for Reproductive Medicine 1995, FIVNAT 1995, Tan 1992, Verlaenen 1995, D'Souza 1997, Bergh 1999). In our study IVF pregnancies 5.5 times more often ended before 37 weeks of gestation, 4 to 6 times more often ended before 32 weeks of gestation and gestational age at delivery was 12 to 14 days shorter than in the control groups. Singleton IVF pregnancies in this study showed to be 1.5 as likely to end before 37 weeks of pregnancy and gestational age at birth was approximately three days shorter than in control pregnancies. The risk of preterm birth < 32 weeks for singleton births was elevated in the comparison with one control group, but not with the other.

In the first years after the introduction of IVF it was believed that IVF singleton pregnancies, once achieved, were as healthy as pregnancies from spontaneous conceptions. Any problems were assumed to be related to the high multiple pregnancies rate in IVF pregnancies. When studies began to appear that showed a negative effect of IVF exposure on the course of the subsequent singleton pregnancies, it was proposed by a number of authors that maternal characteristics instead of procedure related problems, were the cause of the negative outcome (Tan 1992, Verlaenen 1995). In our study, we employed two separate control groups to be able to control for a large number of factors in which the IVF group may differ from the control group and that may influence the risk of preterm birth. The SMOCC-cohort contains pregnancies resulting in liveborn children only and the LVR-cohort contains pregnancies that resulted in antenatal stillbirths as well. In the SMOCC-cohort more data were available on maternal characteristics than in the LVR-cohort, although the SMOCC-data did not contain the data on induction of labour that we needed to correct for iatrogenic preterm birth. Still, the results of the analyses in both groups were similar. Stepwise regression analyses showed that taking into account multiplicity and maternal demographic variables such as age and parity, indeed lowered the relative risk of preterm birth in the IVF group and diminished the difference in duration of pregnancy in days. A second step taking into account other factors such as smoking and alcohol use, pre-existent chronic diseases and whether or not the delivery was induced, appeared to lower the relative risk a little more. Apparently, differences exist between the IVF- and the control groups that may explain part of the extra risk that exists in the IVF group. However, after controlling for many different maternal characteristics such as socio economic status, ethnicity, smoking, and preexistent chronic conditions, IVF mothers still had a 1.2 and 1.6 times higher risk of premature delivery and a duration of pregnancy that was more than 2 days shorter than in the control groups. The risk at preterm birth < 32 weeks was, after correction, no longer elevated.

The effects of IVF were studied in a restricted subsample of low-risk women to avoid that imperfectly measured covariates may result in a spurious association between IVF exposure and preterm birth. Although the numbers in these analyses were necessarily lower than those used in the multivariate models, the effect estimates still clearly show an effect of IVF exposure in all three groups. This lends further credibility to the IVF effect found in our multivariate analyses.

Iatrogenic prematurity does not appear to be an explanation for the increased risk of preterm birth, since we controlled for inductions and primary caesarean sections. More inductions and caesarean sections are being performed in the IVF group than in the control group, but the majority of them take place after 37 weeks of gestation. Idiopathic infertility also, does not seem to explain the added risk of preterm birth in singleton IVF pregnancies. Evidently, the best way to test for the effect of infertility on the outcome of IVF pregnancies would be to compare IVF pregnancies to a group of spontaneous pregnancies after a period of subfertility. Comparing the effect of IVF on the pregnancy outcomes under study, for the four different outcome categories, offers an approximation. The results of this analysis indicate that women with tubal pathology have a slightly higher risk than those with other types of fertility problems. If infertility iself (and the chronic conditions possibly associated with it) contributes to the excess negative outcomes in the IVF cohort, one would expect to find the highest risk in the subgroup of idiopathic or mixed infertility. If a difference in risk truly exists, it appears in our data to be related to tubal pathology instead of to infertility of unknown causes. This makes it more unlikely that some maternal factor associated with infertility may explain the negative effects of IVF.

The validity of our results partly depend on the correct calculation of gestational age. Unfortunately, no gold standard exists with respect to assessment of gestational age, neither in naturally conceived pregnancies, nor in IVF pregnancies. However, since adding 17 days to date of birth minus date of embryo transfer in our calculations, is more likely to over- than underestimate gestational age in the IVF group, our main risk estimates should be conservative. Using 15 and 16 days instead, increased the risk of preterm birth in IVF exposed relative to that in the control groups, but dit not alter our conclusions. If we assume the difference in gestational age between IVF- and controlgroups is overestimated, recalculation will decrease the percentage of IVF children who are small for gestional age (SGA). In another publication from the same database (Koudstaal 2000), using the '17 days' method to calculate duration of pregnancy, it has been shown that IVF children, after correction for maternal factors, are more prone to be SGA. The combination of increased rate of preterm birth and increased risk of SGA indicates that perinatal outcome in IVF pregnancies is worse than in naturally conceived control pregnancies.

If the IVF procedure itself, instead of the mothers' profile leads to an increased risk of adverse perinatal outcomes, a number of different explanations can be offered. Some studies have linked superovulation to reduced birthweight and growth retardation in humans. Elevation of insulin-like growth factor binding protein (IGFBP-1) is the suggested cause (Lino 1986, Howell 1989, Wang 1991, Johnson MR, Abbas 1993, Johnson 1995) which would imply that the ovulation induction within the IVF procedure may be harmful to foetal development. Another study suggests that a higher incidence of abnormal placental shapes exists in IVF singleton pregnancies (22% vs 6% in the control group) and that abnormal umbilical cord insertions are more often found (Jauniaux 1990). The phenomenon may be related to inadequate orientation of the blastocyst after IVF embryo transfer which may cause inferior placental functioning. Ovarian hyperstimulation has been reported to increase the circulating relaxin levels (Johnson 1991), while high serum relaxin concentrations have been found to be correllated to preterm birth (Weiss 1993) and to number of follicles in the

preceding treatment cycle (Kristianson 1996). Olivennes et al (1993) found no difference in preterm birth rate between a group of 162 IVF-patients and 263 infertile controls treated with ovarian hyperstimulation. In cattle and sheep assisted reproduction, culture, embryo micromanipulation and transfer, are known to lead to larger sized offspring, which in these species is a sign of pathology (Wennerholm 1998, Walker 1992, Wagtendonk 2000). The animals donating the gametes, nor the recipients, suffer from infertility. Therefore, these adverse effects, if real, have to be related to aspects of the procedure itself. In animal In Vitro Production studies these adverse effects seem at least partly related to the ovulation induction (Wagtendonk 2000).

One may argue that an increased risk of 1.2 to 1.6 of a preterm birth and 2.3 days difference in duration of pregnancy in IVF singletons is not large. However, it is an indication that the procedure itself may carry risks to the offspring, the long-term effects of which are yet unknown. Further, for pregnancies already compromised and at risk for preterm birth and/or low birthweight an additional risk may carry significant dangers to the wellbeing of the newborn. Studies in severely low birthweight or preterm children have shown that small differences in birthweight or length of gestation may significantly influence (long term) development (Aylward 1989, Stewart 1999).

For individual IVF mothers and clinicians, the outcomes of the uncorrected analysis may be of importance. They should realise that a singleton IVF pregnancy is approximately 1.5 times as likely to result in a preterm delivery than a non-IVF singleton pregnancy. In clinics that transfer more than one embryo per cycle, women should be informed that their risk of a preterm delivery is 5.5 times higher than normal with a subsequent higher risk of mortality and morbidity for the offspring. This finding may have important implications for policy makers as IVF continues to be practised on a wide scale. With the arrival of IVF/ICSI for severe male factor problems, its use is still expanding

Our results are in accordance with those found previously. They clearly show IVF singleton pregnancies are more likely to end prematurely and suggest this excess risk may be related to aspects of the procedure itself. More research needs to be carried out into procedure-related aspects of IVF, to pinpoint a more specific cause of the apparent increase in adverse outcomes after IVF. This may, in turn, lead to improvements in the procedure. Follow-up studies into the health and development of IVF children should provide information about possible problems later in life. The ultimate goal of clinicians, couples and policy makers alike, is to create through IVF the possibility to have a healthy child for subfertile people. In IVF research efforts should be geared not only towards improving the success rate of the procedure, but also towards ensuring an outcome that is as healthy as possible.

References

Anthony S, van der Pal KM, Graafmans WC et al. The reliability of perinatal and neonatal mortality rates: differential underreporting in linked professional registers versus Dutch national statistics. Submitted to the American J of Public Health.

AIHW National Perinatal Statistics Unit. IVF and GIFT Pregnancies in Australia and New Zealand 1990: Fertility Society of Australia. Sydney, AIHW National Perinatal Statistics Unit 1992;1-64.

American Society for Reproductive Medicine/Society for Assisted Reproductive Technology Registry. Assisted reproductive technology in the United States and Canada. Fertility and Sterility 1995;64:13-21.

Australian In-Vitro Fertilization Collaborative Group. In vitro fertilization pregnancies in Autralia and New Zealand, 1979-1985. Medical Journal of Australia 1988;148:432-6.

Aylward GP, Pfeiffer SI, Wright A et all. Outcome studies of low birth weight infants published in the last decade: A metaanalysis. J Pediatrics 1989;115:515-20.

Beral V, Doyle P, Tan SL, Mason BA, Campbell S. Outcome of pregnancies resulting from assisted conception. British Medical Bulletin 1990;46:753-68.

Bergh T, Ericson A, Hillensjö T, Nugren KG, Wennerholm UB. Deliveries and children born after in-vitro-fertilisation in Sweden 1982-95: a retrospective cohort study. The Lancet 1999;354:1579-85.

Biggers JD. In vitro fertilisation and embryo transfer in human beings. N Engl J Med 1981;304:36.

Botting BJ, Davies IM, Macfarlane AJ. Recent trends in the incidence of multiple biths and associated mortality. Archives of Disease in Childhood 1987;62:941-50.

Dhont M, Neubourg de F, Elst van der J, Sutter de P. Perinatal Outcome of Pregnancies After Assisted Reproduction: A Case Control Study. J Assist Reprod Gen 1997;10(14):575-80.

Doyle P, Beral V, Maconochie N. Preterm delivery, low birthweight and small-for-gestational-age in liveborn singleton babies resulting from in vitro fertilization. Human Reproduction 1992;7:425-8.

D'Souza SW, Rivlin E, Cadman J, et al. Children conceived by in vitro fertilisation after fresh embryo transfer. Archives of Disease in Childhood 1997;76:F70-F74.

Fiedler K, Würfel W, Krüsmann G, Rothenaicher M, Hirsch P, Krusmann sen. W. Zum Schwangerschafts- und Geburtsverlauf nach In-Vitro-Fertilisation (IVF). Z Geburtsh u Perinat 1990;194:8-12.

FIVNAT. Pregnancies and births resulting from in vitro fertilization: French national registry, analysis of data 1986 to 1990. Fertil Steril; 1995;64:746-56.

Ghazi HA, Spielberger C, Källén B. Delivery outcome after infertility – a registry study. Fertil Steril 1991;55:726-32.

Gissler M, Malin Silverio M, Hemminki E. In vitro fertilisation pregnancies and perinatal health in Finland 1991-1993. Human Reproduction 1995;10:1856-61.

Goldenberg RL, Cliver SP. Small for gestational Age and intrauterine growth restriction: definitions and standards. Clin. Obstet. Gynaecol 1997;40;704-14.

Herngreen WP, Reerink JD, van Noord-Zaadstra BM, Verloove-Vanhorick SP, Ruys JH. SMOCC: Design of a respresentative cohort-study of live-born infants in the Netherlands. Europ J Public Health 1992;2;117-22.

Howe RS, Sayegh RA, Durinzi KL, Tureck RW. Perinatal Outcome of Singleton Pregnancies Conceived by In Vitro Fertilization: A Controlled Study. J Perinatol 1990;10(3):261-6.

Howell RJS, Economides D, Teisner B, Farkas AG & Chard T. Placental proteins 12 and 14 in pre-eclampsia. Acta Obstetrica et Gynecologica Scandinavica, 1989;68;237-40.

Jauniaux E, Englert Y, Vanesse M, Hiden M & Wilkin P. Pathologic features of placentas from singleton pregnancies obtained by in vitro fertilization and embryo transfer. Obstetric and Gynecology 1990;76;61-4.

Johnson MR, Okokon E, Collins WP, Sharma V, Lightman SL. The effect of human chorionic

gonadotropin and pregnancy on the circulating level of relaxin. J Clin Endocr Metabol 1991;72:1042-7.

Johnson MR, Abbas A, Norman-Taylor JQ et al. Circulating placental protein 14 (PP14) in the first trimester of spontaneous and IVF pregnancies. Human Reproduction 1993;8;323-6.

Johnson MR, Irvine R, Hills F, et al. Superovulation IGFBP-1 and birthweight. European Journal of Obstetrics and Gynaecology and Reproductive Biology 1995;59;193-5.

Kiely JL, Susser M. Preterm birth, intrauterine growth retardation, and perinatal mortality. American Journal of Public Health 1992;82:343-5.

Kloosterman GJ. On intrauterine growth. Int J Gynaecol Obstet 1970;8:895-912.

Koudstaal J, Buitendijk SE, Anthony S, Land JA, Naaktgeboren N, Vermeiden JPW, Verloove-Vanhorick SP, Visser GHA. Neonatal birthweight in singleton pregnancies following in-vitro fertilization. Results of a Dutch multicentre prospective study. 1999. PhD-thesis, University of Utrecht.

Koudstaal 2000. Human Reprod, in press.

Kristiansson P, Svärdsudd, Schoultz B von, Wramsby H. Supraphysiological serum relaxin concentration during pregnancy achieved by in-vitro fertilization is strongly correlated to the number of growing follicles in the treatment cycle. Hum Reprod 1996:11:2036-40.

Lino K, Sjoberg J & Seppala M. Elevated circulating levels of a decidual protein, placental protein 12, in preeclampsia. Obstetric and Gynecology 1986;68;58-60.

McFaul PB, Patel N, Mills J. An audit of the obstetric outcome of 148 consecutive pregnancies form assisted conception: implications for neonatal services. Br J Obstet Gynaecol 1993;100:820-72.

Olivennes F, Rufat P, André B, Pourade A, Quiros MC, Frydman R. The increased risk of complications observed in singleton pregnancies resulting from in vitro fertilization (IVF) does not seem to be related to the IVF method itself. Hum Reprod 1993;8(8):1297-1300.

Petersen K, Hornnes PJ, Ellingsen S, et al. Perinatal outcome after in vitro fertilisation. Acta Obstetricia et Gynecologica Scandinavica 1995;74:129-31.

Reubinoff BE, Samueloff A, Ben-Haim M, Friedler S, Schenker JG, Lewin A. Is the obstetric outcome of in vitro fertilized singleton gestations different from natural ones? A controlled study. Fertility and Sterility 1997;67:1077-83.

Roest J, Verhoef AA, Lent M van, Huisman GJ, Zeilmaker GH. Results of decentralized in vitro fertilization treatment with transport and satellite clinics. Hum Reprod 1995;10:563-7.

Rufat P, Olivennes F, Mouzon de J, Dehan M, Frydman R. Task force report on the outcome of pregnancies and children conceived by in vitro fertilization. Fertility and Sterility 1994;61:324-30.

Saunders DM, Mathews M, Lancaster P. The Australian Register: Current Research and Future Role. Ann Acad Sci 1988;541:7-21.

Stewart AL, Rifkin L, Amess PN et al. Brain structure and neurocognitive and behavioural function in adolescents who were born very preterm. Lancet 1999;353:1653-7.

Tan SL, Doyle P, Campbell S, et al. Obstetric outcome of in vitro fertilization pregnancies compared with normally conceived pregnancies. American Journal of Obstetrics and Gynecology 1992;167:778-84.

Tanbo T, Dale PO, Lunde O, Moe N, Åbyholm T. Obstetric Outcome in Singleton Pregnancies After Assisted Reproduction. Obstet Gynaecol 1995;86:188-92.

Verlaenen H, Cammu H, Derde MP, Amy JJ. Singleton Pregnancy After In Vitro Fertilization: Expectations and Outcome. Obstetrics and Gynecology 1995;86:906-10.

Wagtendonk van-de Leeuw AM, Mullaart E, Roos de APW, Merton JS, Daas den JHG, Kemp B, Ruigh de L. Effects of different reproduction techniques: AI, MOET or IVP, on health and welfare of bovine offspring. Theriogenology 2000; 53:575-597.

Walker SK, Heard TM & Seamark RF. In vitro culture of sheep embryos without co-culture: Successes and perspectives. Theriogenology 1992;37;111-26.

Wang HS, Lim J, English J, Irvine L & Chard T. The concentration of insulinlike growth factor

(IGF-1) and insulin-like growth factor-binding protein-1 (IGFBP-1) in human ubilical cord serum at delivery: Reaction to fetal weight. Journal of Endocrinology 1991;129:459-64.

Wang XJ, Clark AM, Kirby CA, Philipson G, Petrucco O, Anderson G and Matthews CD. The obstetric outcome of singleton pregnancies following in vitro fertilization/gamete intra-Fallopian transfer. Hum Reprod 1994;9(1):141-6.

Weiss G, Goldsmith LT, Sachdev R, Hagen S von, Lederer K. Elevated first-trimester serum relaxin concentrations in pregnant women following ovarian stimulation predict prematurity risk and preterm delivery. Obstet Gynecol 1993;82:821-8.

Wennerholm UB, Albertsson-Wikland K, Bergh C et al. Postnatal growth and health in children born after cryopreservation as embryos. Lancet 1998;351:1085-7.

Wennerholm UB, Janson PO. Wennergren M, Kjellmer I. Pregnancy complications and short-term follow-up of infants born after in vitro fertilization and embryo transfer (IVF/ET). Acta Obstet Gynecol Scand 1991;70:565-573.

Williams MA, Goldman MB, Mittendorf R, Monson RR. Subfertility and the risk of low birthweight. Fertil Steril 1991;56:668-71.

Chapter 4

Birthweight in singleton pregnancies following in vitro fertilization. Results of a Dutch multicentre prospective study

J. Koudstaal[1], S.E. Buitendijk[2], S. Anthony[2], J.A. Land[3], N. Naaktgeboren[4], R. Schats[5], S.P. Verloove-Vanhorick[2] and G.H.A. Visser[1]

[1] Department of Obstetrics and Gynaecology, University Medical Centre, PO Box 85500, 3508 GA Utrecht, the Netherlands.
[2] TNO Prevention and Health, Division of Child Health, PO Box 2215, 2301 CE Leiden, the Netherlands.
[3] Department of Obstetrics and Gynaecology, Academic Hospital Maastricht, PO Box 5800, 6202 AZ Maastricht, the Netherlands
[4] Department of Gynaecology and Reproductive Medicine, Leiden University Medical Centre, PO Box 9600, 2300 RC Leiden, the Netherlands.
[5] IVF Centre, Division of Endocrinology and Fertility, Institute of Endocrinology, Reproduction and Metabolism. Vrije Universiteit Medical Centre, de Boelelaan 117, 1081 HV Amsterdam, the Netherlands.

[1] To whom correspondence should be addressed.

Submitted for publication

Abstract

Conflicting results with regard to neonatal birthweight have been reported in singleton pregnancies after in vitro fertilization. Increased rates of children with low birthweight (LBW) and children that are small-for-gestational age (SGA) have been found by some, although these findings were not confirmed in other studies. We performed a prospective study in which 88% of Dutch IVF clinics participated. The study group consisted of 1465 liveborn IVF singletons. A group of 2061 liveborn non-IVF exposed singletons served as controls. The crude difference in birthweight between the IVF and the control newborns was 186 grams. After adjusting for maternal age and height, ethnic origin, education, parity, previous abortion, pre-existing diabetes and hypertension, smoking habit, alcohol consumption, infant sex, and infant gestational age at birth, birthweight was 90 grams lower in the IVF group than in the controls ($p < 0.001$). The similarly adjusted odds of SGA was 1.4 in the IVF group compared to the control group (95% CI [1.1 - 1.8]). Stratifying by cause of infertility revealed that infant birthweight did not differ significantly between different categories.

Introduction

In 1985, the Australian In-Vitro Fertilization Collaborative Group was the first to mention a high incidence of preterm birth and low birthweight (LBW) in singleton IVF pregnancies. Other studies from different countries subsequently showed similar results (MRC Working Party 1990, Rizk 1991, AIHW National Perinatal Statistics Unit 1992, Friedler 1992, FIVNAT 1995, von During 1995, Gissler 1995). In these studies the outcome of IVF pregnancies was compared to that of the general parturient population. However, IVF mothers on average are older and of lower parity than the general parturient population, which may (in part) explain the less favourable results in IVF pregnancies (Jonas 1991). Several authors (Tan 1992, Tanbo 1995, Verlaenen1995, Dhont 1997, Reubinoff 1997) compared IVF singleton pregnancies to a control group of spontaneously conceived singleton pregnancies matched by age, parity, and a varying number of other factors. In some of these studies (Tan 1992, Tanbo 1995, Verlaenen 1995) an increased rate of preterm birth and LBW in IVF pregnancies was found while others (Dhont 1997, Reubinoff 1997) noted no differences. In a number of studies (Doyle 1992, Olivennes 1993, FIVNAT 1995) comparison with a reference curve showed that the small-for-gestational age (SGA) rate in IVF children significantly exceeded the expected 10%, whereas in other studies (Australian In-Vitro fertilization Collaborative Group 1985 and 1988, Verlaenen 1995, Reubinoff 1997) the SGA rate was not increased. Wang (1994) carried out a regression analysis to evaluate the effect of maternal age and parity, gestational age, cause of infertility, and pregnancy complications (pregnancy induced hypertension and antepartum haemorrhage) on neonatal outcome in 465 IVF pregnancies and a large control population. In the IVF group the proportion of SGA was significantly increased. Bergh et al (1999) studied pregnancy outcomes in over 3000 singleton pregnancies after IVF, stratifying for parity, maternal age and duration of infertility. They reported a 1.5 times increased risk of birthweight under 1500 grams and a 1.2 times increased risk of birthweight under 2500 grams. Both outcomes were

statistically significant. Since they did not correct for gestational age it cannot be derived from their data whether the risk of SGA is larger in IVF than in control pregnancies.

The aim of our study was to compare infant birthweight in singleton IVF pregnancies to infant birthweight in a control population of singletons after adjustment for a number of covariates. We performed a prospective study in which all Dutch IVF clinics were invited to participate. Pregnancies from a well documented parturient population served as the control group. In both groups, only live born singletons were included. The IVF group comprised 1448 pregnancies and the control group 2061 pregnancies. To our knowledge, this is the first study in which population size and available information are sufficient to enable controlling for a substantial number of covariates simultaneously.

Methods

The study was carried out by TNO Prevention and Health in cooperation with the IVF Working Group of the Dutch Society of Obstetrics and Gynaecology. The Netherlands has 12 IVF centres, 8 University and 4 non-University affiliated. The IVF centres perform the complete IVF procedure for their own patients. Most centres further participate in the IVF program of a number of other hospitals usually called 'transport hospitals' (Jansen 1986). Transport hospitals execute all steps in the IVF procedure, except for the laboratory phase and the embryo transfer, which is performed by the IVF centre. Some transport hospitals collaborate with other hospitals called 'satellite hospitals' (Roest 1995). In the latter case three hospitals participate in the treatment cycle. The harvesting of the oocytes is performed by the transport hospital, the laboratory phase and embryo transfer by the IVF centre, while the remaining parts of the program i.e., patient selection, ovarian hyperstimulation, cycle monitoring and follow-up after embryo transfer are taken care of by the satellite hospital. Altogether 12 IVF centres, 23 transport hospitals, and 6 satellite hospitals participate in the Dutch IVF program. All hospitals except one transport hospital agreed to participate in the study. The study protocol was approved by the ethical committee of TNO, and by the ethical committees of all IVF centres.

IVF group

The study was conducted in the period from January 1, 1994 up until December 31, 1995. If in these two years an IVF procedure resulted in a pregnancy, diagnosed primarily by ultrasound, the couple received written information about the study as well as an informed consent form from their IVF clinic. In case they chose to participate the signed form was returned to the research team. Data on fertility history and IVF treatment were obtained through the participating IVF clinics. Women who subsequently miscarried before 16 weeks of pregnancy were excluded from the study. Two months after the expected date of delivery the patient received a questionnaire containing questions about her obstetric and medical history as well as about the course and outcome of the pregnancy. The provider of the obstetrical care (general practitioner, midwife or gynaecologist) was also asked through a questionnaire for information concerning

pregnancy and delivery. For this study, only live born IVF singletons were analysed as the control group comprised live born children only.

Control group

Singleton pregnancies from the Social Medical Survey of Children attending Child Health Clinics (SMOCC) study served as controls (see Herngreen et al., 1992 for a detailed description of the SMOCC study.) In the SMOCC study data were collected through postpartum interview with all mothers from live born infants in the areas of 21 different child health clinics. Data were gathered concerning maternal socio-demographic and health characteristics, the course of pregnancy, and health of the newborn on 2127 children (2061 singletons), born between April 1st 1988 and October 31st 1989.

Non-participation

During the data collection period, not all IVF clinics that originally agreed to participate, were able to obtain consent from all eligible women in their centre. Of the 41 IVF clinics, five clinics (four transport and one satellite clinic) contributed less than 5% of their potential to the study and were excluded. Of the remaining 36 clinics, 11 (one IVF centre and ten transport hospitals) obtained consent from less than 80% of their eligible patients while 25 clinics obtained informed consent from 80% to 100% of their pregnant IVF patients. Inquiries as to the reasons for non-participation revealed that under-participation was generally due to temporary logistic or staff problems such as pregnancy- or sick leave, or departure of the physician coordinating the data collection. Judging from the reasons for non-participation, it seems unlikely that it has introduced a bias for the following reasons: firstly, the IVF clinics in the Netherlands usually serve their own region and patient selection is not based on risk profile. It is therefore unlikely that the non-participating clinics had patient populations with different pregnancy outcomes than the participating centres. Secondly, in clinics that participated partially, logistic and staff problems, rather than patient refusal, appeared to have led to incomplete data collection. Overall, patients proved eager to participate. The few refusals that did occur were mostly from immigrants and were related to language problems.

Definition of main variables

Gestational age at delivery in IVF pregnancies was defined as the difference between birth date and the day of embryo transfer with 17 days added. In most studies as an approximation of the first day of last menstrual period (LMP) used for calculating gestational age in IVF pregnancies, the date of oocyte puncture is taken and 14 days added, or the day of embryo transfer is taken and 16 days added. However, since embryo transfer is either performed 2 or 3 days after puncture, we chose to use the date of embryo transfer and add 17 days. Gestational age in SMOCC pregnancies was defined as the difference between the date of birth and the date of LMP. Preterm delivery was defined as delivery before 37 completed weeks and a delivery before 32 completed weeks as very preterm delivery. Small-for-gestational age (SGA) was defined as birthweight below the 10th percentile of the national reference curve and very-small-for gestational age (VSGA) as birthweight below the 2.3rd percentile. The Dutch reference curve is corrected for gestational age, parity and sex of the infant (Kloosterman, 1970). Low birthweight (LBW) was defined as birthweight < 2500 g and very low birthweight (VLBW) as birthweight < 1500 g. Digit preference was assessed for birthweight. It

showed to occur only for 5 and 10 gram groups and in a similar fashion in IVF- and SMOCC-groups.

Statistical analysis
The χ^2-test was used to compare categorical outcome variables between the IVF and SMOCC group. Continuous variables were compared by analysis of variance (ANOVA). Linear regression analysis was carried out in order to adjust the difference in birthweight between IVF and SMOCC pregnancies for maternal and pregnancy related factors. Factors were entered in a stepwise fashion. The factors taken into account were maternal factors (age, height, education, ethnic origin, parity, previous abortions, pre-existing diabetes and hypertension), lifestyle factors (smoking habit, alcohol consumption), infant sex, and gestational age at birth. The crude and adjusted differences are presented with 95% confidence intervals (CI)].

Similarly, logistic regression analysis was carried out to calculate adjusted odds ratio for the risk of SGA and LBW. A separate analysis was carried out to compare the risk of low birthweight in subgroups of women with tubal pathology, male factor infertility, idiopathic infertility and infertility of mixed causes.

Lastly, analyses were carried out comparing IVF- and control pregnancies after restricting both the IVF- and the control groups to low-risk subgroups homogeneous on all factors but the IVF exposure. A subgroup was constructed consisting of singleton pregnancies in women pregnant for the second time, who were between the ages of 20 and 35, were not belonging to an ethnic minority, were non-smoking, non-drinking, and had no pre-existent diabetes or hypertension. Thus, if covariates were measured imperfectly, they are less likely at this low level of risk and in these 'pure' subsamples, to alter the effect of IVF exposure on the outcomes measured and result in a spurious association between IVF exposure and low birthweight.

In order to ensure that non-participation did not bias the results of our study, the data were analysed in two ways. Firstly, the entire data set, including patient data from 'incomplete' centres, was analysed. Secondly, the analysis was repeated with data only from the 25 centres that provided information on more than 80% of their pregnant IVF patients. If similar outcomes were found, it may be assumed that our results were not biased by non-participation.

Results

The IVF group comprised 1465 singleton pregnancies and the SMOCC group 2061 singletons. The maternal characteristics of the IVF and control group are summarized in table 1. IVF mothers were older, of lower parity, and more frequently had a history of spontaneous abortion. They also were, on average, higher educated. Small differences were found in maternal height and ethnic origin. IVF mothers consumed less alcohol during their pregnancy than control mothers. No difference was found in the percentage

of smokers, nor in the average number of cigarettes smoked. The rate of pre-pregnancy hypertension was comparable. Diabetes mellitus was reported less in the IVF group.

Table 1 Characteristics of women in the IVF- and SMOCC control cohort

	IVF (n = 1465) % (n)	SMOCC# (n = 2061) % (n)	p-value
Average maternal age ± SD*	33.7 ± 3.8	28.9 ± 4.5	< 0.001
Maternal age (yrs)			
< 25	0.6 (9)	16.2 (334)	< 0.001
25-29	13.3 (194)	40.3 (831)	
30-34	43.2 (628)	32.7 (674)	
35-39	36.5 (530)	9.4 (194)	
≥ 40	6.4 (93)	1.1 (23)	
Average maternal height (cm) ± SD*	168.9 ± 6.8	168.3 ± 6.9	0.02
Ethnicity			
Dutch	93.3 (1327)	91.2 (1847)	0.03
non-Dutch	6.7 (96)	8.8 (179)	
Education			
below university degree	71.9 (989)	83.4 (1657)	< 0.001
university degree	28.1 (386)	16.6 (330)	
Parity			
primiparous	68.6 (1004)	42.6 (863)	< 0.001
multiparous	31.4 (460)	57.4 (1164)	
Previous miscarriages			
yes	26.4 (386)	17.9 (363)	< 0.001
no	73.6 (1078)	82.1 (1664)	
Pre-existent diabetes			
yes	0.3 (5)	1.1 (22)	0.01
no	99.7 (1444)	98.9 (2005)	
Pre-existent hypertension			
yes	1.5 (22)	2.2 (45)	0.14
no	98.5 (1426)	97.8 (1982)	
Smoking during pregnancy			
no smoking	76.0 (1070)	74.0 (1498)	0.18
≥ 1 cigarette/day	24.0 (338)	26.0 (527)	
Alcohol use during pregnancy			
no alcohol use	83.4 (1173)	73.9 (1497)	< 0.001
≥ 1 glass/week	16.6 (234)	26.1 (528)	

liveborn infants from the Social Medical Survey on Children Attending Child Health Clinics
* SD = Standard Deviation
Totals may vary because of missing values

Pregnancy and neonatal outcome of the IVF and SMOCC pregnancies are summarized in table 2. The mean gestational age at birth in IVF pregnancies was 2.8 days shorter and significantly more pregnancies in this group ended preterm. The rate of very preterm birth was also higher in the IVF group although this difference did not reach statistical significance. The rates of LBW, VLBW, SGA and VSGA were higher in the IVF group.

After logistic regression analysis, adjusting for maternal age, maternal height, parity, previous abortion, pre-existing diabetes and hypertension, ethnic origin, education, smoking, alcohol consumption, infant sex, and gestational age at birth the SGA risk was significantly higher for IVF children whereas the adjusted LBW risk was not. For VLBW and VSGA, regression analysis was not applicable since the numbers were too small. The crude and adjusted odd ratios are shown in table 3.

Table 2 Condition of children in the IVF- and control cohort

	IVF (n = 1465) % (n)	SMOCC (n = 2061) % (n)	p-value
Sex			
- boys	52.3 (761)	49.2 (1014)	0.07
- girls	47.7 (693)	50.8 (1047)	
Average birthweight in grammes ± SD*	3247 ± 628	3433 ± 546	< 0.001
Birthweight (grammes)			
< 1500	1.3 (19)	0.4 (9)	< 0.001
1500 – 1999	2.8 (4.1)	0.9 (18)	
2000 – 2499	5.9 (86)	3.3 (69)	
2500 – 2999	19.2 (278)	13.1 (269)	
3000 – 3499	35.2 (509)	32.4 (667)	
3500 – 3999	26.1 (378)	35.7 (736)	
≥ 4000	9.5 (137)	14.2 (292)	
Growth < P10 (SGA)			
- yes	13.8 (199)	9.7 (196)	< 0.001
- no	86.2 (1244)	90.3 (1831)	
Growth < P2.3 (VSGA)			
- yes	3.6 (52)	1.7 (34)	< 0.001
- no	96.4 (1391)	98.3 (1993)	
Gestational age at birth (days) ± SD*	275.6 ± 15	278.4 ± 13	< 0.001
Preterm delivery (< 32 wks)			
yes	1.1 (16)	0.5 (11)	0.06
no	98.9 (1433)	99.5 (2049)	
Preterm delivery (< 37 wks)			< 0.001
yes	9.3 (135)	5.2 (108)	
no	90.7 (1314)	94.8 (1952)	

* SD = Standard Deviation

Table 3 Initial and adjusted odds ratios for birthweight characteristics for IVF versus control pregnancies.

Birthweight Characteristics	Crude odds ratio [95% CI[2]]	Adjusted[1] Odds ratio [95% CI[2]]	Adjusted[1] p-value
Birthweight < 2500 g	2.2 [1.7-2.9]	1.3 [0.8-2.1]	0.22
Birthweight < 1500 g	3.0 [1.4-6.7]	*	
Birthweight < 10th percentile	1.5 [1.2-1.9]	1.4 [1.1-1.8]	0.02
Birthweight < 2.3rd percentile	2.2 [1.4-3.4]	*	

[1] After adjustment for maternal factors (age, height, ethnic origin, education, parity, previous abortions, pre-existing diabetes and hypertension), lifestyle factors (smoking, alcohol consumption), infant sex, and gestational age at birth

[2] 95% Confidence Interval

* Regression analysis not applicable because of small numbers

Figure 1 shows a histogram of the birthweight percentiles of IVF and SMOCC children demonstrating a shift to the left for the IVF group ($p < 0.001$). Table 4 shows stepwise logistic analysis with mean birthweight as outcome. The crude mean birthweight of IVF children is 186 g lower than that of control children After linear regression analysis, adjusting for maternal demographic and lifestyle factors, infant sex and gestational age, the remaining difference was 90 g.

Fig 1 Birthweight percentiles in IVF and SMOCC pregnancies

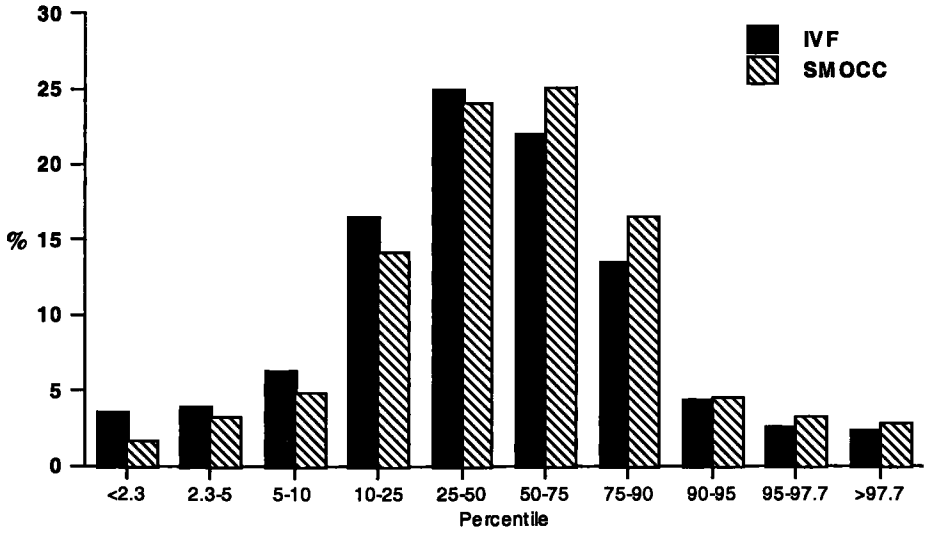


Table 4 Stepwise linear regression analysis on the difference in birthweight in grams between the IVF and control group

Factors entered	Mean birthweight [95% Confidence Interval] IVF v.s. SMOCC
None	-186 [-226 to -145]
Maternal pre-pregnancy factors (age, height, ethnicity, education, parity, previous abortions, pre-existing diabetes and hypertension)	-162 [-209 to -114]
Maternal pre-pregnancy plus lifestyle factors (smoking, alchohol consumption)	-147 [-195 to -99]
Maternal pre-pregnancy plus lifestyle factors plus infant sex and gestational age	-90 [-127 to -52]

Mean birthweight, mean gestational age at birth, and rates of preterm birth, LBW and SGA stratified by the cause of infertility are summarized in table 5.

*Table 5 Perinatal outcome of IVF pregnancies of live-born singletons stratified by cause of infertility**

Outcome	Tubal (n=360)	Idiopathic (n=254)	Male factor (n=272)	Other (n=451)	p-value
Gestational age at birth (days)**	275 ± 16	275 ± 14	276 ±13	276 ± 15	0.31
Preterm delivery < 37 weeks (%)	11.8	9.5	8.5	8.5	0.38
Birthweight (g)** ± SD	3249 ± 657	3216 ± 668	3265 ± 599	3245 ± 611	0.85
Birthweight < 2500 g	10.7	10.3	9.9	9.7	0.97
Birthweight < 10th percentile	14.4	14.3	12.5	12.9	0.86

* Of 128 women the cause of infertility was not stated
**Values are means ± SD

The IVF- and SMOCC-cohorts were compared on risk of low birthweight and SGA in a restricted low-risk subsample containing 260 SMOCC-women and 88 IVF women. The risk estimate for LBW was 3.7 (95% CI [1.1 to 12.5]) and for SGA 2.1 (95% CI [0.9 to 4.9]). Difference in birthweight was -176 grams (95% CI [-310 to -44 grams]). We corrected for maternal age, height and education in the restricted sample. This yielded an odds ratio for LBW of 5.4 (95% CI [1.2 – 25.6]) and for SGA of 2.9 (95% CI [0.97 to 9.1]) and a difference in birthweight of –61 grams (95% CI[-184 to 61]).

In table 6 results are shown from a logistic regression analysis comparing the risk of low birthweight between the different indication categories while controlling for maternal and life style factors and gestational age. Although there is an indication that the risk of low birthweight is higher in the group of sperm abnormalities, idiopathic infertility and the combination group than in group of tubal pathology, this difference does not reach

statistical significance. Type of infertility is not clearly a factor influencing the low birthweight risk. A similar regression analysis with mean birthweight as outcome variable also shows no statistically significant differences between the indication groups (results not shown).

Table 6 Logistic regression analysis comparing the risk of low birthweight between the different indication groups within the IVF-cohort of live-born singletons

	n (%)	Odds ratio [95% CI]# Birthweight < 2500 grs		P-value
Indication *				
tubal pathology	328 (26.6)	1.0		
sperm abnormalities	251 (20.4)	2.0	[0.9 – 4.4]	0.1
idiopathic infertility	234 (19.0)	1.5	[0.7 – 3.6]	0.3
combination	418 (34.0)	1.9	[0.9 – 3.9]	0.1

: 95% Confidence Interval

* : controlling for age, maternal height, parity, ethnicity, mother's education, previous abortions, smoking, alcohol use, pre-existent hypertension, pre-existent diabetes, infant sex and duration of pregnancy

Analyses were repeated for the subgroup only of participating centres that contributed over 80% of their patients to our study. The results show decimal point changes in risk estimates only. We therefore conclude that our results were not biased by non-participation.

Discussion

The crude mean birthweight in 1465 IVF children was 186 grams lower than in the group of 2061 control singletons ($p < 0.001$). After adjustment by linear regression analysis for a maternal prepregnancy and lifestyle factors, infant sex and gestational age, mean birthweight of IVF singletons was 90 g lower than that of control singletons ($p < 0.001$). This lower birthweight was also demonstrated by an 1.4 times increased SGA rate ($p=0.02$). Furthermore, IVF children were over-represented in the lower percentile groups and under-represented in the higher percentile groups (Fig1, $p < 0.001$). Mean gestational age at birth was 2.8 days lower in the IVF group ($p < 0.001$). The preterm birth rate in the IVF group was increased in comparison to the control group (9.3% vs 5.2%, $p < 0.001$).

These results presume correct calculation of gestational age, both in IVF and SMOCC pregnancies. However, gestational dating is not an exact science (Goldenberg and Cliver, 1997). In control pregnancies gestational age was calculated from the first day of the LMP. In IVF pregnancies gestational age was calculated from an artificial LMP that was created by subtracting 17 days from the date of embryo transfer. The difference in gestational age between IVF- and control pregnancies might be under- as well as overestimated. If we assume that the difference in gestational age is overestimated, then the difference in preterm rate between the two groups may in reality be smaller. However, the difference in adjusted birthweight and SGA rate will be larger. Conversely, underestimation of the difference in gestational age will result in a decrease in difference in SGA rate, but in a larger difference in preterm birth rate (Rufat 1994).

In conclusion: the combination of increased preterm and SGA rates in IVF pregnancies indicates that perinatal outcome in these pregnancies is less favourable than in control pregnancies. One may argue that a difference in mean birthweight of 90 g is very small and not of clinical significance. This may be true for the individual, although early growth impairment may increase health risks in later life (Barker, 1993). Furthermore, a significantly larger number of infants is of very low birthweight which poses considerable health risks in itself. Last, the decrease in mean birthweight may be an indication of adverse effects of the IVF procedure per se on the development of the foetus.

In order to confirm the results from our regression models, the effects of IVF were studied in a restricted subsample of low-risk women. Although the numbers in these analyses are necessarily much lower than those used in the multivariate models and the confidence intervals around the point estimates wider, the effect estimates still point towards an effect of IVF exposure. This further supports the IVF effect found in our multivariate analyses.

A history of infertility has been linked by some authors to an increased risk of LBW and SGA (Ghazi 1991, Williams 1991), whereas others found no such relation (Tuck 1988, Li 1991). Patients with infertility of tubal origin differ from those with infertility of other causes. Since most tubal infertility is acquired, these women would probably be fertile had their tubes been functioning normally, while couples with unexplained or male factor infertility may represent a population with inherent impaired reproductive potential. Thus, if infertility *per se* causes the adverse neonatal outcome, one would expect that this effect was most pronounced in cases with unexplained or male infertility. According to Wang (1994) the risk for SGA is indeed largely confined to women with unexplained infertility, whereas Doyle (1992) found a suggestion of an increased risk of LBW in unexplained or male infertility. Tanbo (1995) found no relation between the cause of infertility and neonatal outcome. Although our data are suggestive of a lower risk of low birthweight in the group of women with tubal pathology, this difference did not reach statistical significance. We have to conclude that no clear relationship seems to exist in our study between type of infertility and mean birthweight or risk of low birthweight, indicating that the adverse neonatal outcomes cannot be explained by the type of infertility. Further studies are needed to elucidate the reasons for the difference in growth and birthweight between IVF- and control children. Our data do suggest that apart from maternal factors and the increased risk of preterm birth, certain aspect of the IVF procedure itself negatively influence the birthweight of IVF babies. If indeed the IVF procedure itself causes and increased risk of adeverse outcomes, it has yet to be determined which aspects of the procedure may influence this risk.

References

AIHW National Perinatal Statistics Unit. Assisted conception Australia and New Zealand 1990: Fertility Society of Australia. Sydney. AIHW National Perinatal Statistics Unit 1992;1-64.

Australian In-Vitro Fertilization Collaborative Group. High incidence of preterm births and early losses in pregnancy after in vitro fertilization. Br Med J 1985;291:1160-3.

Australian In-Vitro Fertilization Collaborative Group. In-vitro fertilization pregnancies in Australia and New Zealand, 1979-1985: Med J Australia 1988;148:429-36.

Barker DJP. The intrauterine origins of cardiovascular disease. Acta Paediatr Suppl 1993;391:93-9.

Bergh T, Ericson A, Hillensjö T, Nugren KG, Wennerholm UB. Deliveries and children born after in-vitro-fertilisation in Sweden 1982-95: a retrospective cohort study. The Lancet 1999;354:1579-85.

Dhont M, Neubourg de F, Elst van der J, Sutter de P. Perinatal Outcome of Pregnancies After Assisted Reproduction: A Case-Control Study. J Assist Reprod Gen 1997;10:575-80.

Doyle P, Beral V, Maconochie N. Preterm delivery, low birthweight and small-for-gestational-age in liveborn singleton babies resulting from in-vitro fertilization. Hum Reprod 1992;7:425-8.

Düring, von V, Forsdahl JMMF, Åbyholm T, Kolvik R, Ertzeid G, Steier A, Irgens VBLM. Svangerskap, fødsel og barn etter in vitro-fertilisering i Norge, 1988-91. Særtrykk av Tidsskrift for Den norske lægeforening. Tidsskr Nor Lægeforen 1995;115:2054-60.

FIVNAT. Pregnancies and births resulting from in vitro fertilization: French national registry, analysis of data 1986 to 1990. Fertil Steril 1995;64:746-56.

Friedler S, Mashiach S, Laufer N. Births in Israël resulting from in-vitro fertilization/embryo transfer, 1982-1989: National Registry of the Israëli Assoation for Fertility Research. Hum Reprod 1992;7:1159-63.

Ghazi HA, Spielberger C, Källén B. Delivery outcome after infertility - a registry study. Fertil Steril 1991;55:726-32.

Gissler M, Malin Silverio M, Hemminki E. In-vitro fertilization pregnancies and perinatal health in Finland 1991-1993. Hum Reprod 1995;10:1856-61.

Goldenberg R.L, Cliver S.P. Small for gestational age and intrauterine growth restriction: definitions and standards. Clin Obstet Gynaecol 1997;40:704-14.

Herngreen WP, Reerink, JD, van Noord-Zaadstra BM, Verloove-Vanhorick SP, Ruys JH. SMOCC: Design of a representative cohort-study of live-born infants in the Netherlands. Europ J Public Health 1992;2:117-22.

Jansen CAM, Beek JJ van, Verhoeff A, Alberda AT, Zeilmaker GH. In vitro fertilization and embryo transfwer with transport of oocytes. Lancet 1986;I:676.

Jonas O, Chan A, Roder D, Macharper T. Pregnancy outcomes in primigravid women aged 35 years and over in South Australia, 1986-1988. Med J Australia 1991;154:246-9.

Kloosterman GJ. On intrauterine growth. Int J Gyneacol Obstet 1970;8:895-912.

Li TC, Macleod I, Singhal V, Duncan SLB. The obstetric and neonatal outcome of pregnancy in women with a previous history of infertility: a prospective study. Br J Obstet Gynecol 1991;98:1087-92.

MRC Working Party on Children Conceived by In Vitro Fertilisation. Births in Great Britain resulting from assisted conception, 1978-87. Br Med J 1990;300:1229-33.

Olivennes F, Rufat P, André B, Pourade A, Quiros MC, Frydman R. The increased risk of complication observed in singleton pregnancies resulting from in-vitro fertilization (IVF) does not seem to be related to the IVF method itself. Hum Reprod 1993;8:1297-1300.

Reubinoff BE, Samueloff A, Ben-Haim M, Friedler S, Schenker JG, Lewin A. Is the obstetrical outcome of in vitro fertilized singleton gestations different from natural ? A controlled study. Fertil Steril 1997;6:1077-83.

Rizk B, Doyle P, Tan SL, Rainsbury P, Betts J, Brinsden P, Edwards R. Perinatal outcome and congenital malformations in in-vitro fertilization babies from the Bourn-Hallam group. Hum Reprod 1991;6:1259-64.
Roest J, Verhoeff A, Lent M van, Huisman GJ, Zeilmaker GH. Results of decentralized in-vitro fertilization treatment with transport and satellite clinics. Hum Reprod 1995;10:563-7.
Rufat P, Olivennes F, de Mouzon J, Dehan M, Frydman R. Task force report on the outcome of pregnancies and children conceived by in vitro fertilization (France: 1987 to 1989). Fertil Steril 1994;61:324-30.
Tan SL, Doyle P, Campbell S, Beral V, Rizk B, Brinsden, Mason B, Edwards RG. Obstetric outcome of in vitro fertilization pregnancies compared with normally conceived pregnancies. Am J Obstet Gynaecol 1992;167:778-84.
Tanbo T, Dale PO, Lunde O, Moe N, Åbyholm T. Obstetric Outcome in Singleton Pregnancies After Assisted Reproduction. Obstet Gynaecol 1995;86:188-92.
Tuck SM, Yudkin PL, Turnbull AC. Pregnancy outcome in elderly primigravidae with and without a history of infertility. Br J Obstet Gynaecol 1988;95:230-7.
Verlaenen H, Cammu H, Derde MP, Amy JJ. Singleton Pregnancy After In Vitro Fertilization: Expectations and Outcome. Obstet Gynaecol 1995;86:906-10.
Wang XJ, Clark AM, Kirby CA, Philipson G, Petrucco O, Anderson G and Matthews CD. The obstetric outcome of singleton pregnancies following in-vitro fertilization/gamete intra-Fallopian transfer. Hum Reprod 1994;9:141-6.
Williams MA, Goldman MB, Mittendorf R, Monson RR. Subfertility and the risk of low birth weight. Fertil Steril 1991;56:668-6.

Chapter 5

Higher birthweight in IVF pregnancies after embryo-cryopreservation

S.E.Buitendijk[1], S. Anthony[1], D.D.M. Braat[2], F.M. Helmerhorst[3], A.H.M. Simons[4], J. Koudstaal[5], P.H. Verkerk[1]

[1] TNO Prevention and Health, Division of child health, PO Box 2215, 2301 CE Leiden, the Netherlands.
[2] Department of Obstetrics and Gynaecology, University Hospital St. Radboud, PO Box 9101, 6500 HB Nijmegen, the Netherlands.
[3] Department of Gynaecology and Reproductive Medicine, Leiden University Medical Centre, PO Box 9600, 2300 RC Leiden, the Netherlands.
[4] Department of Obstetrics and Gynaecology, Academic Hospital Groningen, PO Box 30001, 9700 RB Groningen, the Netherlands.
[5] Department of Obstetrics and Gynaecology, University Medical Centre, PO Box 8500, 3508 GA Utrecht, the Netherlands.

Submitted for publication

Abstract

Pregnancy outcomes of 2460 IVF-children were studied, 99 of whom were born from embryo's that were previously frozen and subsequently thawed and 2361 after cycles with fresh embryo transfer. Multiple pregnancy rate as well as maternal age and height were statistically significantly lower in the "cryo' than in the "fresh" group. The groups did not differ in the other characteristics. Children born after cryopreservation had half the risk of low birthweight (OR=0.5, 95%CI [0.3-0.8]) and weighed 234 grams more (95% CI [92-377]) than children born after fresh embryo transfer, after correcting for the difference in multiple pregnancy rate and maternal age and height. The difference in risk of growth retardation below the tenth percentile between the groups just ceased to be statistically significant after correction (OR=0.5, 95% CI [0.3-1.04]). No difference in gestational age could be observed between the groups after correction. In 89% of cryopregnancies no hyperstimulation was used. In the pregnancies after fresh embryo transfer, no hyperstimulation was used in 4% of preceding cycles. The results indicate that children born after cryopreservation have better perinatal outcomes than those born after IVF with transfer of fresh embryo's.

Introduction

The superovulation method as employed in most In Vitro Fertilisation (IVF-) treatments regularly yields a high number of oocytes to be fertilised and, thus, a high number of embryo's. Often, only a limited number of embryo's is being transferred into the uterus to prevent large multiple pregnancies. In most IVF-clinics as part of the regular treatment programme, the surplus embryo's are being frozen. They are thawed in subsequent cycles. The first successful pregnancy after cryopreservation was described in 1983 (Trounson) and the first birth after cryopreservation was reported from the Netherlands in 1994 (Zeilmaker). Cryopreservation and thawing involve major cellular changes and it is as yet not clear whether they may have adverse effects on the course of the pregnancy. A small number of studies into the outcome of pregnancies after cryopreservation have been reported (Wada 1994, Heijnsbroek 1995, Suttcliffe 1995, Olivennes 1996, Wennerholm 1997, 1998). Heijnsbroek (1995) reported outcomes of 30 cryopregnancies. No control group was employed. Singleton cryo-infants more often than expected were born in breech presentation and had higher average birthweights than the Dutch norm population. Sutcliffe (1995) reported a statistically significant decrease in birthweight and gestational age in 91 IVF pregnancies after cryopreservation compared to a group of normally conceived children. However, it is known from previous studies that IVF pregnancies are more likely to end prematurely and result in a low birthweight child than regular pregnancies (Tan 1992, Gissler 1995). Olivennes (1996) reported prevalence of low birthweight and small for gestational age in 82 IVF-children after cryopreservation without comparing them to a control group. In the two remaining studies (Wada 1994, Wennerholm 1997) 283 and 270 IVF-children born after cryopreservation were studied. No differences were reported in prevalence of preterm birth, mean birthweight or prevalence of low birthweight between singleton cryo-children and control children born after IVF with fresh embryo transfer. Wada (1994) reported a decrease in low birthweight prevalence in cryo-twins, compared to twins in the regular IVF-group.

Recently, as part of a large retrospective cohort study into the course and outcome of IVF pregnancies (Bergh 1999) comparisons of preterm birth rate and risk of low birthweight between 359 children born after IVF with frozen embryo's and 3762 born after conventional IVF, were reported. Children after cryopregnancies were half as likely to be born prematurely and to be of low birthweight than children after conventional IVF.

Aytoz et al (1999) studied obstetric outcomes in cryopregnancies after IVF and ICSI compared to fresh IVF- and ICSI-pregnancies. They reported no differences in singleton pregnancies in preterm birth or low birthweight risk, or in preterm birth rate in twin pregnancies. In twin pregnancies, however, the risk of low birthweight was twice as high in "fresh" pregnancies than in cryopregnancies.

Development of children after cryopreservation during the first 2 years of life has in existing studies not been reported to be disturbed (Sutcliffe 1995, Olivennes 1996, Wennerholm 1998).

The total number of cryopreserved pregnancies studied to date is small. Although the results appear to be reassuring, no firm conclusions can be drawn as yet. Animal data are contradictory with respect to the consequences of freezing and thawing on the development of embryo's. A number of mouse studies indicate that severely damaged cryopreserved mouse embryo's are lost early in pregnancy and cannot negatively influence perinatal outcome (Ashwood-Smith 1985, Mobraaten 1987, Sidyakina 1991). However, a recent study found differences in morphophysiological and behavioural features between cryopreserved and control mice (Dulioust 1995). The relevance of these data for the human situation remains to be determined. The aim of the present study was to contribute to the existing knowledge by describing perinatal outcomes in a cohort of IVF pregnancies after cryopreservation, compared to similar IVF pregnancies after fresh embryo transfer.

Material and methods

The study was carried out by TNO Prevention and Health in collaboration with the IVF Working Group of the Dutch Society of Obstetricians and Gynaecologists, the 12 Dutch IVF-centres holding laboratory facilities, plus the centers collaborating with them.

From January 1st 1994, to December 31st 1995, women who were interested in participating were asked written informed consent. It was obtained shortly after the IVF-treatment had shown to result in an ongoing pregnancy, diagnosed primarily with fetal heart action at ultrasound. Women who subsequently miscarried before 16 weeks of pregnancy, were excluded from the study. Data on infertility history, treatment schedule, stimulation protocol, and cryopreservation, were obtained through the participating IVF-clinics. Two months after the expected date of delivery the IVF-patient received a mailed questionnaire containing questions about her obstetric and medical history, the course and outcome of her pregnancy as well as demographic characteristics.

The provider of the prenatal care (general practitioner, midwife or obstetrician) was asked to provide information concerning the prenatal period and delivery, through a questionnaire. Methods are described in more detail elsewhere (Buitendijk 2000, in preparation).

Statistics

Differences in number of embryo's transferred and in maternal and pregnancy characteristics between the children born after cryopreservation and after fresh embryo transfer, were described and tested for statistical significance. Differences in risk of low birthweight (birthweight < 2500 grams), preterm birth (birth < 37 weeks) and Small for Gestational Age (SGA) (growth < p10) as well as differences in mean birthweight and mean gestational age, were calculated and tested for statistical significance. Pearson's Chi-square test was used for categorical variables and Student's t-test for continuous variables. Two-tailed tests were employed. Logistic regression models were used to test for the effect of cryopreservation on risk of low birthweight, preterm birth and growth retardation, while controlling for other factors, such as multiplicity of the birth. Linear regression analysis was used to test the effect of cryopreservation on average birthweight and gestational age, while taking these other factors into account.

Results

The total cohort consisted of 2460 IVF-children, 99 of which were born after cryopreservation cycles and 2361 after cycles with fresh embryo transfer. Table 1 shows multiple pregnancy rates and maternal characteristics of the children in both cohorts. Multiple pregnancy rate as well as maternal age and height were statistically significantly lower in the "cryo' than in the "fresh" group. The groups did not differ in the other characteristics shown in the table. Differences in number of embryo's transferred are shown in table 2. More often in the "cryo" than in the "fresh" group three embryo's were transferred (49.5% vs. 39.5%). Differences in ratio of children born and number of embryo's transferred between the "cryo" and the "fresh" group are shown in table 3. It shows, for instance, that in 34.2% of the "cryo" pregnancy group, three embryo's had to be transferred for one child born, while this was the case in 24.0% of the "fresh"group. Table 4 and 5 show the risk of adverse pregnancy outcomes in the total group of children after cryopreservation versus those after fresh embryo transfer both before and after correcting for multiple pregnancy and maternal height and age. Children after cryopreservation have a lower risk of birthweight < 2500 grams and small for gestational age below p10 than those in the group of fresh embryotransfer. They have higher average birthweight and average gestational age. The preterm birth rate is not different between the two groups. After correcting for the difference in multiple pregnancy rate and maternal age and height, the difference in risk of low birthweight remained significant. Children born after cryopreservation have half the risk of low birthweight (OR=0.47, 95%CI [0.26-0.86]) and weigh 234 grams more (95% CI [92-377]) than children born after fresh embryo transfer. The difference in risk of growth retardation < P10 between the groups just ceased to be statistically significant (OR=0.66, 95%CI [0.28-1.04])). No difference in gestational age could be observed between the groups

after correction. In 63 of 67 cryopregnancies, information was available about the use of hyperstimulation. In 56 of those pregnancies (89%) no hyperstimulation was used. In the pregnancies after fresh embryo transfer, no hyperstimulation was used in 4% of preceding cycles.

Table 1 Characteristics of pregnancies after cryopreservation and after fresh embryo transfer

	"cryo" (n = 99) % (n)	"fresh" (n = 2361) % (n)	p-value
Multiplicity			
singleton	68 (67)	57 (1328)	0.03*
part of multiple	32 (32)	43 (1023)	
Maternal age (yrs)	32.7	33.6	0.02*
Maternal height (cm)	168	169	0.04*
Parity			
primiparous	73 (72)	70 (1646)	0.5
multiparous	27 (27)	30 (715)	
Mothers ethnicity			
Dutch	90 (87)	93 (2133)	0.1
non-Dutch	10 (10)	7 (152)	
Maternal education			
below university	81(78)	72 (2003)	0.2
university degree	19(18)	28 (354)	
Maternal smoking			
yes (≥ 1 cig / day)	25 (24)	22 (492)	0.8
no	75 (73)	78 (1768)	
Maternal alcohol use			
yes	17 (16)	15 (344)	0.7
no	83 (81)	85 (1911)	

* statistically significant

Table 2 Number of embryo's transferred in pregnancies with cryopreservation and with fresh embryo transfer

Number of embryo's transferred	"cryo" (n = 76) % (n)	"fresh" (n = 1797) % (n)
1	11.8 (9)	4.4 (79)
2	40.8 (31)	56.5 (1016)
3	46.1 (35)	36.6 (658)
4	1.3 (1)	2.4 (44)

Table 3 Ratio of number of embryo's transferred to number of children born, in "fresh" and "cryo" pregnancies

Ratio	"cryo"(n= 76) % (n)	"fresh" (n= 1787) % (n)
4 : 1	1.3 (1)	1.7 (30)
3 : 1	34.2 (26)	24.0 (429)
2 : 1	32.9 (25)	43.8 (783)
2 : 2	9.2 (7)	10.2 (182)
4 : 3	-	0.1 (1)
1 : 1	22.4 (17)	19.9 (356)
2 : 3	-	0.2 (5)
1 : 2	-	0.1 (1)

Table 4 Effect of cryopreservation on risk of low birthweight (LBW) < 2500 grams, SGA < P10 and preterm birth < 37 weeks

Variables in model	Odds ratio [95% CI] * LBW < 2500 gr	Odds ratio [95% CI] * SGA < P10	Odds ratio [95% CI] * birth < 37 wks
Cryo-preservation	0.5 [0.3 – 0.8]	0.4 [0.3 – 0.7]	0.8 [0.5 – 1.3]
Cryo-preservation, multiplicity, maternal age and maternal height	0.5[0.3 – 0.9]	0.5 [0.3 – 1.04]	0.9 [0.4 – 1.5]

* 95% Confidence Interval

Table 5 Effect of cryopreservation on birthweight and duration of pregnancy

Variables in model	Grams difference in birthweight [95% CI] *	Days difference in gestational age [95% CI] *
Cryo-preservation	+338 [163 – 513]	+6 [2 – 11]
Cryo-preservation, multiplicity, maternal age and maternal height	+234[92 – 377]	0 [-2 to +2]

* 95% Confidence Interval

Discussion

Our results indicate that children born after cryopreservation have better perinatal outcomes than those born after IVF with transfer of fresh embryo's. The fact they are less often part of a multiple pregnancy lowers their risk of adverse outcomes. However, even after correction for multiplicity, their risk of low birthweight is half the risk of the fresh transfer group and their average birthweight is over 200 grams more. Although the comparison of risk of SGA just failed to reach statistical significance, the results are suggestive of a lowered risk for children after cryopreservation.

These results are not contradicting the results of previous studies. Studies that compare cryo-

singletons to control singletons all suggest better, rather than worse outcomes (Wada 1994, Heijnsbroek 1995, Olivennes 1996, Wennerholm 1997) although these differences do not reach statistical significance. Bergh (1999) reports a significantly higher birthweight and lower risk of preterm birth in the subgroup of cryopregnancies, corrected for maternal age and parity, but not for multiplicity of pregnancy. Wada (1994) reported significantly higher birthweights in cryo-twins compared to twins after fresh embryo-transfer. In our analysis we corrected for those variables in which the cryo-group differed from the regular IVF-group. It is possible that in the previous studies not showing statistical significance, the cryo-mothers' demographic characteristics predisposed them to negative outcomes. If this was indeed the case, not correcting for them may have underestimated a truly existing positive effect of cryopreservation on perinatal outcomes.

If, indeed, cryopregnancies have better perinatal outcomes than regular IVF pregnancies do, a number of possible explanations can be offered. One explanation may be that the hyperstimulation in the IVF-procedure adversely affects the health of the embryo and the course of the pregnancy. A related explanation may be that the endometrium in unstimulated cycles is of better quality than in stimulated cycles, which may positively influence the growth of the embryo. In general, IVF-singleton pregnancies (the majority of which are from stimulated cycles) are more likely to result in low birthweight and growth retarded babies than non-IVF control pregnancies (Tan 1992, Gissler 1995, Koudstaal 1999). Most cryo-IVF pregnancies in our study (89%) resulted from unstimulated cycles, while in only 4% of regular IVF-cycles no hyperstimulation was used. Similar percentages were reported in the previous studies (Wennerholm 1997). Olivennes (1993) reported higher risks of low birthweight in a non-IVF-group after stimulation than in a group of spontaneously conceived pregnancies. Some studies have linked elevation of insulin like growth factor binding protein (IGFBP-1) that occurs with superovulation to reduced birthweight and growth retardation (Johnson 1993, Johnson 1995). Yet another explanation may be that the higher percentage of cryo-embryo's lost during the pregnancy, results in a 'survival of the fittest'. The embryo's chosen to be preserved by freezing in most IVF-clinics are of lesser quality than the ones transferred immediately and loss during pregnancy of cryo-embryo's is larger than that of freshly transferred embryo's. Partially damaged embryo's have indeed been shown to have lower viability than fully intact embryo's (Van den Abbeel, 1997). However, this process may result in a natural selection of only very healthy cryo-embryo's surviving the early period of development.

Whatever the explanation, from the present data it may be concluded that cryopreservation of IVF-embryo's does not lead to an increase in adverse perinatal outcomes. It may even yield better outcomes than regular IVF-cycles. However, more research should be carried out into the course and outcome of IVF pregnancies in unstimulated cycles. At present, the potential negative effects of cryopreservation cannot be distinguished from the potential positive effects of using unstimulated cycles in IVF-treatments. Further, although the present data are reassuring, more information is still needed on the long-term effects and the development of children after cryopreservation, since some developmental disturbances can only be detected later in life.

References

Abbeel E van den, Camus M, Waesberghe L van, Devroey P, Steirteghem AC van. Viability of partially damaged human embryo's after cryopreservation. Hum Reprod 1997;12(9):2006-10.

Ashwood-Smith MJ. Genetic damage is not produced by normal cryopreservation procedures involving either glycerol or dimethylsulfoxide. A cautionary note, however, on possible effects of dimethyl sulfoxide. Cryobiology 1985;22:427-33.

Aytoz A, Abbeel E van den, Bonduelle M, Camus M, Joris H, Steirteghem AC van, Devroey P. Obstetric outcome of pregnancies after the transfer of cryopreserved and fresh embryos obtained by conventional in-vitro fertilization and intracytoplasmic sperm injection. Hum Reprod 1999;14(10):2619-24.

Bergh T, Ericson A, Hillensjö T, Nygren KG, Wennerholm UB. Deliveries and children born after in-vitro fertilisation in Sweden 1982-95: a retrospective cohort study. The Lancet 1999;354:1579-85.

Buitendijk SE, Koudstaal J, Anthony S, Leertveld RA, Hogerzeil HV, Jansen K. Preterm birth in pregnancies following in-vitro fertilization. Results of a Dutch multicentre prospective study. 2000 In preparation.

Dehan M and Frydman R. Perinatal outcome and follow-up of 82 children aged 1-9 years old conceived from cryopreserved embryo's. Hum Reprod 1996;11(7):1565-8.

Dulioust E, Toyama K, Busnel M, et al. Long-term effects of embryo freezing in mice. Proceedings of the National Academy of Sciences of the United States of America, 1995;92:589-93.

Gissler M, Malin Silverio M, Hemminki E. In vitro fertilisation pregnancies and perinatal health in Finland 1991-1993. Human Reproduction 1995;10:1856-61.

Heijnsbroek I, Helmerhorst FM, Berg van den – Helder AF, Zwan LJ van, Naaktgeboren N, Keirse MJNC. Follow-up of 30 pregnancies after embryo cryopreservation. Eur J Obstet Gynecol Reprod Biol 1995;59:201-4.

Johnson MR, Abbas A, Norman-Taylor JQ, et al. Circulating placental protein 14 (PP14) in the first trimester of spontaneous and IVF-pregnancies. Human Reproduction, 1993;8:223-6.

Johnson MR, Irvine R, Hills F, et al. Superovulation IGFBP-1 and birthweight. Eur J Obstetrics and Gynaecology and Reproductive Biology, 1995;59:193-5.

Koudstaal J, Buitendijk SE, Anthony S, Land JA, Naaktgeboren N, Vermeiden JPW, Verloove-Vanhorick SP, Visser GHA. Neonatal birthweight in singleton pregnancies following in-vitro fertilization. Results of a Dutch multicentre prospective study. 1999. PhD-thesis, University of Utrecht.

Mobraaten LE, Bailey DW. Effects of freezing mouse embryo's on mutation rate. Cryobiology 1987;24:586.

Olivennes F, Rufat P, André B, et al. The increased risk of complication observed in singleton pregnancies resulting from in-vitro fertilization (IVF) does not seem to be related to the IVF method itself. Human Reproduction 1993;8:1297-1300.

Olivennes F, Schneider Z, Remy V, Blanchet V, Kerbrat V, Fanchin R, Hazout A, Glissant M, Fernandez H, Sidyakina TM, Gollimbet VE. Viability and genetic stability of the bacterium Escherichia coli HB 101 with the recombinant plasmid during preservation by various methods. Cryobiology 1991;28:251-4.

Sidyakina TM, Golimbet VE. Viability and genetic stability of the bacterium Escherichia coli HB 101 recombinant plasmid during preservation by various methods. Cryobiology 1991;3:251-4.

Sutcliffe AG, D'Souza SW, Cadman J, Richards B, McKinlay IA, Lieberman B. Outcome in children from cryopreserved embryo's. Archives of Diseases in Childhood 1995;72:290-3.

Tan SL, Doyle P, Maconochie N et al. Pregnancy and birth rates of live infants after in vitro fertilization in women with an without previous in vitro fertilization pregnancies: a study of eight thousand cycles at one center. Am J Obstet Gynecol 1994;170:34-40.

Trounson AO and Mohr L. Human pregnancy following cryopreservation, thawing and transfer of

an eight-cell embryo. Nature 1983;305:7.
Wada I, Macnamee MC, Wick K et al. Birth characteristics and perinatal outcome of babies conceived from cryopreserved embryo's. Hum Reprod 1994;9:543-6.
Wennerholm UB, Hamberger L, Nilsson L, Wennergren M, Wikland M and Bergh C. Obstetric and perinatal outcome of children conceived from cryopreserved embryo's. Hum Reprod 1997;12(8):1819-25.
Wennerholm UB, Albertsson-Wikland, Bergh C, Hamberger L, Niklasson A, Nilsson L, Thiringer K, Wennergren M, Wikland M, Borres MP. Postnatal growth and health in children born after cryopreservation as embryo's. Lancet 1998;351:1085-90.
Zeilmaker GH, Alberda AT et al. Two pregnancies following transfer of intact frozen-thawed embryo's. Fertil Steril 1984;42:293-6.

Chapter 6

Development of IVF children and controls at two years of age

S.E. Buitendijk[1], C.A. Dorrepaal[1], A.L. den Ouden[1], P.H. Verkerk[1], R.S.G.M. Bots[2], N.S. Macklon[3], A.H.M. Simons[4], S.P. Verloove-Vanhorick[1]

[1] TNO Prevention and Health, Division of Child Health, PO Box 2215, 2301 CE Leiden, the Netherlands.
[2] Department of Obstetrics and Gynaecology, St. Elisabeth Hospital, PO Box 90151, 5000 LC Tilburg, the Netherlands
[3] Department of Obstetrics and Gynaecology, Academical Hospital Rotterdam, PO Box 2040, 3000 CA Rotterdam, the Netherlands
[4] Department of Obstetrics and Gynaecology, Academical Hospital Groningen, PO Box 30001, 9700 RB Groningen, the Netherlands

Submitted for publication

Abstract

A total of 197 IVF children and 194 control children were examined with the Bayley test. IVF mothers less often than control mothers had completed higher education (34% vs. 45%), they were less likely to be single parents (1% vs. 5%) and were more likely to take full time care of the IVF child (35% vs. 24%). Thirteen IVF children were born before 36 weeks of gestation, 4 of the controls were. Correction was made in the analyses for differences in parity, age of the mother, education of the mother, single parenthood and caretaking situation. The average test scores for IVF- and control children on the mental and motor developmental index (DI) and on their mental and motor developmental age (DA) with and without correction were the same in the IVF- and the control group. Both the IVF- and control group scored higher than the norm. On the mental scale, no differences in K-scores lower than 5 could be observed between the IVF- and the control group. A K-score lower than 5 indicates clinically significant behavioural problems. On the motor K-score 3.0% (n=6) of IVF children had a score lower than 5 compared to 0% in the control group, a statistically significant difference. Half (n=3) of the low-scoring children were born before 36 weeks of pregnancy and the average duration of pregnancy was 3 weeks lower in this subgroup than in IVF children born at term. The results of our analysis suggest that IVF children develop normally and that the negative effect on motor development is caused primarily by the excess of preterm births in the IVF group compared to the controls.

Introduction

The first IVF baby is an UK citizen who celebrated her 21st birthday in 1999. Directly after her birth, she was declared to be 'completely healthy' by 4 paediatricians who, together, performed 41 tests (Edwards and Steptoe 1980). The founders of IVF, embryologist Edwards and gynaecologist Steptoe, stated that this conclusion had not surprised them (Edwards and Steptoe 1980). The apparent normalcy of the first IVF baby, the enthusiasm with which the new technology was embraced, the focus on the success rate of the procedure, or a combination of these factors, may explain why in the early years of IVF little attention was being paid to the condition of the IVF offspring. Only during the late 1980's and early 1990's, studies started to emerge in medical journals on the course and outcome of IVF pregnancies. These studies showed that IVF pregnancies were of shorter gestation and more often resulted in low birthweight children (Australian In-Vitro Fertilization Collaborative Group 1985, Beral 1990, MRC Working Party on Children Conceived by In Vitro Fertilization 1990, Doyle 1992, Rufat 1994, Petersen 1995, D'Souza 1997) than normally conceived pregnancies. This effect largely appeared to be due to the high multiple pregnancy rates after IVF treatment, but partly persisted when only singleton pregnancies were being analysed.

Even in the late 1990's, however, only a limited number of studies have been published that describe the development of IVF babies **after** the direct postnatal period (Mushin 1986, Morin 1989, Brandes 1992, Raoul Duval 1994, Cederblad 1996, Ron-El 1996). Most of these compare the development of IVF children and control children during the first two years of life using either the Bayley Scales of Infant Development or comparable scales.

None show a difference in development between IVF children and the control group.

Half of these studies (Mushin 1986, Raoul Duval 1994, Ron-El 1996) include numbers that are so small (only approximately 30 children in each group) they are probably insufficient to show differences between the IVF and the control group, would these truly exist. The larger studies contain around 100 IVF children each. One study (Cederblad 1996) compared IVF children to a norm population. With this approach effects of the IVF procedure cannot be distinguished from maternal and demographic effects. It is well known that IVF mothers differ from other mothers in a number of characteristics, such as age and parity, that may influence developmental scores. The remaining two studies (Morin 1989, Brandes 1992) compare the IVF group to a control group, matched to the IVF group on a number of variables. Morin (1989) did not match for number of older children in the family and did not distinguish between children born preterm or at term. Brandes (1992) matched pairs of IVF and control children for a large number of variables such as number of older children in the family, parents' socio-economic status, mother's age and preterm birth and low birthweight. The effect of preterm birth of low birthweight could therefore not be separated from the effect of the IVF treatment itself. IVF children more often have a lower gestational age and a lower birthweight (Wang 1994, Gissler 1995, Reubinoff 1997, Buitendijk 2000, Koudstaal 1999). Another factor that has so far not been taken into account, is the quality of IVF parenting. Golombok (1996) compared families with a child conceived by IVF, to families with a naturally conceived child. Parent-child interaction was superior in the IVF families, possibly indicating a higher degree of 'wantedness'. This may well enhance the IVF children's score on developmental tests.

In our study, we included a larger number of children than has been done so far and controlled for a substantial number of factors known or suspected to differ between the IVF group and the general population of pregnant women. We thus hoped to establish whether IVF children, on average, have different psychomotor development than other children. Further, we attempted to differentiate between the effects of IVF exposure and those of maternal and social characteristics on the development of the IVF children.

Materials and methods

TNO Prevention and Health together with the IVF Working Group of the Dutch Society of Obstetricians and Gynaecologists and 12 Dutch IVF centres with laboratory facilities, carried out a large study into the course and outcome of IVF pregnancies. From January 1st 1994 to December 31st 1995 women whose IVF treatment had resulted in an ongoing pregnancy were asked written informed consent to take part in our study. Information about the perinatal period and the delivery as well as the first month of development of the baby were obtained shortly after birth from the mother, the provider of the prenatal care and the paediatrician if the baby had been examined by one. Details of the study are described elsewhere (Buitendijk 2000).

Of the cohort of over 2000 deliveries thus obtained, 200 liveborn children from singleton pregnancies were examined with the Bayley Scales of Infant Development (van der Meulen 1984) at two years of age. The first selection criterion was, for logistic reasons, that the mother should live in the Leiden region. If a woman upon entering into the study had consented for her baby to be examined at two years and if we possessed no information on the baby having died, a letter was sent to ask renewed permission for examination of the child's psychomotor development around the 2nd birthday. IVF mothers differ from mothers in the general population on a number of characteristics, such as age, education and number of previous children that may influence the scores on the children's developmental test. We therefore felt we needed to compare the IVF children to a control group that was matched on these variables as well as to compare them to the Dutch norm population. Control children were obtained from three Child Health Clinics in the Leiden region. Since we expected differences in mother's age and parity between IVF- and control children to be so large that we may have difficulty controlling for them in our analyses, we matched IVF- and control children on those two variables. For each IVF child a control was selected from one of the Child Health Clinics on the basis of the following criteria: 1. the control child was from a singleton pregnancy; 2. the control child matched the IVF child with respect to the mother's age (+/- 6 months); 3. the control and the IVF child had the same number of older siblings. Personnel from the Child Health Clinic approached parents from suitable control children first by telephone and subsequently by letter. All re-approached IVF parents consented. Fewer than 10 intended control parents refused participation, the reason stated in all cases that they were 'too busy'.
Control and IVF parents received a questionnaire, to be filled out before the actual examination of the child. The questionnaire contained questions concerning the pregnancy and the child's health and illnesses. It further contained questions on the parent's educational level, family situation (is the child living with two parents or one), and care taking situation (is the child in day care, is there a caretaker other than the parent in the home). All children were examined in their own home by one of two trained Bayley test technicians between October 1996 and March 1998 at the age of two years +/- two weeks. Parents had been asked beforehand not to reveal the child's IVF- or control status to the test technician. All parents received a written report with the test results. In case the technician suspected abnormal development, she informed the parents and sometimes advised consultation by a paediatrician.

The Bayley Scales of Infant Development measures early mental and motor development in children aged 3 to 30 months (van der Meulen 1984). The scores consist of a sum score on the mental scale, verbal and non-verbal, (range 0 to 163 points) and a sum scores on the motor scale (0 to 81 points). From these two sum scores the following scores can be derived by comparing the sum scores with the age-adjusted norm: 1) K-scores on the mental and motor scale, ranging from 0 to 10 with 6 being the norm; 2) Development Index on the mental and motor scale, ranging from 50 to 150 with 100 being the norm; and 3) test age: if the calculated test age equals the real chronological age of the child, the score is equal to the norm. The K-score and Developmental Index can be calculated from the sum scores both with and without correcting for gestational age. Correcting for a gestational age of 32 weeks

in a 2-year old child implies that the sum score of this child is compared to the norm sum score for children of 22 months instead of the norm sum score for children of 24 months. The K-score can further be divided in scores below or above 5, which is the cut-off point for a clinically significant developmental delay. Preterm birth is in our study defined as birth before 36 completed weeks of gestation, since in the Bayley Scales 36 weeks is used as a cut-off point below which corrections for gestational age can be made.

Statistical methods

Sum- and K-scores, Developmental index and test age were calculated for children in the IVF- and the control group, for both groups as a whole as well as after stratification into a term birth and a preterm birth (defined as < 36 completed weeks) group. Differences in mean scores between the IVF- and the control group were tested for statistical significance using regression analysis with and without correcting for maternal age, number of older siblings, gestational age, sex of the child, maternal education, family situation and care taking situation. The risk of having a K-score < 5 was calculated for the IVF- and control group, and it's difference tested for statistical significance with and without correcting for maternal age, number of older siblings, gestational age, sex of the child, maternal education, family situation and care taking situation, using a logistic regression model. A p-value of < 0.05 was considered statistically significant.

Results

A total of 199 IVF children and 195 control children were visited at home. One IVF mother and 5 control mothers could not be visited for a variety of reasons. Those reasons included an unexpected hospital admission (IVF mother), a move abroad, a last minute cancellation, lack of time because of delivery of a new baby, and a wrongly recorded birth date because of which the child turned out to be too old at the planned test date. Three children (two in the IVF and one in the control group) were not co-operative enough during the test to obtain useful data from them. In total, 197 IVF children and 194 control children could be properly examined with the Bayley test. Table 1 shows the distribution of maternal, pregnancy and caretaking characteristics of the IVF- and control group. IVF mothers less often than control mothers had completed higher education (34% vs. 45%, $p < 0.05$)). They were less likely to be single parents (1% vs. 5%, $p < 0.05$) and were more likely to take fulltime care of the child (35% vs. 24%, $p < 0.05$). Of the IVF children, 7% were born before 36 weeks of gestation, of the control children 2% ($p < 0.05$). In table 2 test scores are shown for IVF- and control children on the mental and motor developmental index (DI) with and without correction for gestational age and on the mental and motor development age (DA). The average scores of IVF children in all domains are lower than those of control children. The differences, however, are very small and none of them are statistically significant. Correcting for variables in which IVF- and control children differ, did not change the statistical significance. As is shown in table 2, both the IVF- and control group score around 8 to 10 points higher than the norm population.

Table 3 and 4 show the results for the subgroups of IVF- and control children that were born at term and preterm, respectively. These also do not differ between the IVF- and the control group. Table 5 shows the number and percentage of children in both groups with a K-score lower than 5, indicating a clinically significant developmental delay. In the IVF group 7.1% (n=14) of children have a K-score on the mental scale of lower than 5, compared to 4.6% (n=9) in the control group. This difference is not statistically significant. On the motor scale, 3.0% (n=6) of the IVF children have a score lower than 5 compared to 0% in the control group. This difference is statistically significant. Table 6 shows the characteristics of the six IVF children with a low K-score < 5 compared to the other IVF children. On average, the low-scoring IVF children had slightly older mothers, were of shorter gestational age and had a lower birthweight as compared to the other IVF children.

Table 1 Distribution of maternal, pregnancy and caretaking characteristics of the IVF- and control group

Characteristics	IVF group (n=199) n (%)	Control group (n = 195) n (%)
Maternal age at birth of child (means ± SD)	34.1 ± 3.6	34.0 ± 3.6
Number of older children in family (means ± SD)	0.2 ± 0.4	0.3 ± 0.6
* Maternal education		
low	64 (32)	40 (21)
medium	63 (32)	65 (33)
high	72 (34)	88 (45)
* Family situation		
two parent family	195 (99)	185 (95)
one parent family	1 (1)	10 (5)
* Caretaking situation		
7 days/wk parental care	70 (35)	47 (24)
5-6 days/wk parental care	66 (33)	59 (30)
< 5 days/wk parental care	62 (32)	87 (46)
* Gestational age < 36 weeks		
no	186 (93)	191 (98)
yes	13 (7)	4 (2)
Sex of child		
boy	92 (46)	104 (53)
girl	107 (54)	91 (47)

Totals may differ because of missing values
* $p < 0.05$ IVF vs. control group

Table 2 Scores Bayley-test all IVF- and Control children

	IVF Children (n = 197)	Control Children (n = 194)	p-value	p-value (after correction) **	Norm Population Score
Developmental Index (DI)					
DI – mental	109.90 ± 17.44	110.60 ± 16.00	0.68	0.52	100
DI – mental (corrected*)	110.36 ± 17.19	110.76 ± 15.92	0.81		100
DI – motor	108.21 ± 14.10	109.70 ± 12.80	0.27	0.32	100
DI – motor (corrected*)	108.55 ± 13.98	109.79 ± 12.77	0.36		100
Developmental Age (DA)					
DA – mental	26.2 ± 2.8	26.4 ± 2.6	0.57	0.45	24
DA – motor	25.4 ± 2.8	25.7 ± 2.5	0.33	0.39	24

* for gestational age

** correction for maternal age, number of older siblings, gestational age, sex of the child, maternal education, family situation and caretaking situation.

Table 3 Scores Bayley-test, term IVF- and control-children

	Term IVF children (n = 184)	Term control Children (n = 190)	p-value	p-value after correction *
Developmental Index (DI)				
DI - mental	111.09 ± 16.43	111.00 ± 15.75	0.86	0.94
DI - motor	108.95 ± 13.02	109.84 ± 12.84	0.58	0.67
Developmental Age (DA)				
DA - mental	26.4 ± 2.7	26.4 ± 2.6	0.75	0.84
DA – motor	25.6 ± 2.6	25.7 ± 2.5	0.66	0.55

* correction for maternal age, number of older siblings, gestational age, sex of the child, maternal education, family situation and caretaking situation.

Table 4 Scores Bayley-test prematurely born IVF- and control-children

	Premature IVF children (n = 13)	Premature Control Children (n = 4)	p-value	p-value after correction *
Developmental Index (DI)				
DI - mental	93.15 ± 22.87	91.75 ± 19.05	0.91	0.67
DI - mental (corrected) **	100.08 ± 24.18	99.50 ± 22.66	0.97	
DI – motor	96.75 ± 23.51	103.25 ± 9.91	0.61	0.89
DI - motor (corrected) **	102.42 ± 24.50	107.25 ± 9.91	0.71	
Developmental Age (DA)				
DA – mental	23.6 ± 3.4	23.3 ± 3.0	0.85	0.67
DA – motor	23.3 ± 4.4	24.3 ± 1.9	0.67	0.86

* correction for maternal age, number of older siblings, gestational age, sex of the child, maternal education, family situation and caretaking situation.

** gestational age corrected.

Table 5 Logistic regression model for K-score < 5 on the Bayley-test for all IVF- and control-children

Outcome parameters Bayley-test	IVF children (n = 197)	Control children (n = 194)	p-value uncorrected	p-value corrected *
K-mental < 5	14 (7.1%)	9 (4.6%)	0.21	0.59
K-motor < 5	6 (3.0%)	0 (0.0%)	0.004	-

* correction for maternal age, number of older siblings, gestational age, sex of the child, maternal education, family situation and caretaking situation. Correction for K-motor < 5 could not be made since zero children in the control group had this score.

Table 6 Comparison of the six IVF children with K-score < 5 to the IVF children with K-score ≥ 5 on a number of characteristics (means ± SD)

Characteristics	IVF K<5 (n = 6)	IVF K≥5 (n = 193)	p-value
Maternal age at birth of child (years) ± SD	3.5 ± 3.5	34 ± 3.6	< 0.001
Gestational age (weeks) ± SD	36.3 ± 4.7	39.2 ± 1.8	< 0.001
Birthweight (grams) ± SD	2006 ± 876	3352 ± 599	0.09

SD = standard deviation

Discussion

Until now, only six studies have been published on the psychomotor development of IVF children around two years of age (Mushin 1986, Morin 1989, Brandes 1992, Raoul Duval 1994, Cederblad 1996, Ron-El 1996). Two of those studies (Mushin 1986, Cederblad 1996) employ no formal control group and instead use the norm population as reference. This approach is problematic since the average norm scores on intelligence tests and developmental tests, such as the Bayley test, have increased over time (Wolke 1994). As a consequence, most general population study groups will now have higher average scores on the Bayley test than the norm average that is mostly derived from data that are at least a decade old. Therefore, the IVF children that in these studies have scores that are 'well within range', may in reality be showing a poorer performance than contemporary norm population children would. Our results show that the IVF- as well as the control group have sum scores that are approximately 10 points higher than the existing Dutch norm score of 100 points. In the Netherlands a study is presently being started that should yield new norm data for the Bayley Scales (Koopmans, written communication).

Our study is the first to employ fairly large numbers of IVF children and controls. The largest so far that did not employ a norm population as control group (Brandes 1992) contained 116 children in both groups, the others had far smaller numbers. The three remaining studies not using norm data (Morin 1989, Raoul-Duval 1994, Ron-El 1996) compared 83, 33 and 30 IVF children respectively to similar numbers of controls. Only large differences in development between IVF- and control children can be detected using these numbers. In fact, until now, only the study by Brandes (1992) was large enough and employed a proper control group to yield reliable conclusions.

In our study, a large number of maternal demographic and pregnancy variables were measured both in the IVF- and in the control group. It is well known that in general, parents of IVF children are older, have not previously had any children and are higher educated than parents in the general population. Parental age, education as well as the presence of older children in the family may influence the developmental outcome in children. Further, sex, gestational age and singleton status, are factors that may influence the Bayley score. By measuring these variables and controlling for them we were in the analysis able to distinguish between an effect of the IVF procedure itself and an effect of maternal, family and pregnancy variables on the development of the IVF children. Furthermore we were able to study the effect of different factors on the developmental outcome. The only two factors on which we matched the two groups at the start of the study were maternal age and presence of older children in the family. The effect of these two variables on developmental outcome could therefore not be studied separately in the IVF- and control group. Our results show that with correction only for maternal age and parity as well as after correction for a large number of maternal and pregnancy variables, IVF children have average developmental scores that are not different from those in control children. Further, we found that controlling in the analysis only for variables in which IVF mothers 'positively' differ from control mothers (family situation and whether or not the mother was the child's sole care taker) did not change the effect estimates. The suggestion from other studies that IVF children may perform better because IVF mothers have better parenting skills and a more stable family situation (Golombok 1996) could not be confirmed by our data. Analysing our data with and without correction for maternal education, maternal age, family and caretaking situation did not influence the results. It is therefore unlikely that in our study a possible negative IVF effect is camouflaged by the more positive family situation of IVF parents. In the study of Brandes (1992) maternal education was not controlled for, although parity and maternal age were. In this study also, no difference was found in average scores between the IVF- and control group. None of the studies performed to date found differences in mean developmental scores between IVF- and control children. In our study also, no differences in mean scores were found between the IVF- and the control group. The IVF group did have lower mean scores in all domains, but these differences were very small and nowhere reached statistical significance. However, when using the clinically significant cut-off point of K-score < 5, it was shown that more IVF children than control children had a risk of having a developmental delay in motor function. Since in the subgroup of IVF children with low motor K-scores the average gestational was three weeks lower than in the rest of the IVF children and 50% of these children were born before 36 weeks compared to 10% in the rest of the IVF group, this difference can probably be largely ascribed to the higher prevalence of preterm births in the IVF group. None of the six previously reported studies report the risk of low K-scores in the IVF- and control groups. Brandes et al (1992) in their study corrected for a large number of variables (birthweight, gestational age, multiplicity, birth order, and mode of delivery, sex, age of the child and age of the mother) before data collection. Even if they had employed the K-score as outcome variable, a deleterious effect of IVF on development of the child through preterm birth and/or growth could not be detected.

Our results need confirmation from other large studies. If not a chance finding, our results suggest that as a group IVF singletons may be more likely to display abnormal motor

development than their normally conceived peers. This effect is not exhibited through a lowering of mean scores, but only through an increase in the percentage of children with a problem score and only in the motor skills. If indeed a real difference in risk of abnormal motor development exists between IVF- and control children, it appears to be strongly related to the shorter gestational age and lower birthweight in the IVF group. A number of studies of IVF pregnancies have reported an excess risk of prematurity in singleton IVF children (Tan 1992, Gissler 1995, Reubinoff 1997, Buitendijk 2000). The risk of a birth before 37 weeks of pregnancy appears to be approximately two times higher in IVF singletons than in controls.

Our data show that, overall, two-year old IVF children have good scores on the Bayley test. In general, the results do suggest that parents of term IVF babies need not worry that IVF exposure per se may negatively influence their child's development. Policy makers and clinicians however, should be aware that the higher risk of short gestation in IVF babies may have consequences for the children's development. Further, many developmental problems, especially those in the cognitive and social-emotional behaviour areas, will not be detected until the child has reached the age of approximately 5 years (Aylward 1989, McCormick 1997, Stewart 1999). More subtle differences in development can only be detected with increased maturity of the brain. In order to determine whether the difference in development that was shown in our data is 'real' or merely a chance finding, or whether differences exist that can be detected only at an older age, the children should be re-examined when they are older. IVF parents are well aware of the special circumstances under which their child was conceived and are eager to receive balanced information on their child's longer-term development. They have also shown to be willing to cooperate with long term research. At the moment, there is reason for reassurance, since the vast majority of IVF children perform well within the normal range. However, our results suggest a certain degree of vigilance is in order and that longer term follow-up should be performed before we may conclude with more confidence whether or not IVF children are at an increased risk of developmental delay.

References

Australian In-Vitro Fertilization Collaborative Group. High incidence of preterm births and early losses in pregnancy after in vitro fertilization. BMJ (Clinical Research ed.) 1985;291:1160-3.
Aylward GP, Pfeiffer SI, Wright A. Outcome studies of low birthweight infants published in the last decade: A metaanalysis. J Pediatrics 1989;115:515-20.
Beral V, Doyle P, Tan SL, Mason BA, Campbell S. Outcome of pregnancies resulting from assisted conception. British Medical Bulletin 1990;46:753-68.
Brandes JM, Scher A, Izkovics J, et al. Growth and development of children conceived by in vitro fertilization. Pediatrics, 1992;90:424-9.
Buitendijk SE, Koudstaal J, Anthony S, et al. Preterm birth in pregnancies following in vitro fertilization. Results of a Dutch multicentre prospective study. In preparation.
Cederblad M, Friberg B, Ploman F, et al. Intelligence and behaviour in children born after in-vitro fertilization treatment. Human Reproduction, 1996;11:2052-7.
Doyle P, Beral V, Maconochie N. Preterm delivery, low birthweight and small-for-gestational-age in liveborn singleton babies resulting from in-vitro fertilization. Human Reproduction, 1992;7:425-8.

D'Souza SW, Rivlin E, Cadman J, et al. Children conceived by in vitro fertilisation after fresh embryo transfer. Archives of Disease in Childhood, 1997;6:F70-F74.
Edwards R, Steptoe P. A matter of life. 1980; © MCMLXXX Finestride Ltd. and Crownchime Ltd.
Gissler M, Malin Silverio M, Hemminki E. In vitro fertilisation pregnancies and perinatal health in Finland 1991-1993. Human Reproduction 1995;10:1856-61.
Golombok S, Brewaeys A, Cook R, et al. The European study of assisted reproduction families: family functioning and child development. Human Reproduction, 1996;11:2324-31.
Koudstaal J, Buitendijk SE, Anthony S, et al. Birth weight in singleton pregnancies following in vitro fertilization. Results of a Dutch multicentre prospective study. Thesis, University of Utrecht, 1999.
McCormick MC. The outcome of very low birthweight infants: Are we asking the right questions? (Commentaries) Pediatrics 1997;99:869-76.
Morin NC, Wirth FH, Johnson DH, et al. Congenital malformations and psychosocial development in children conceived by in vitro fertilization. Journal of Pediatrics, 1989;15:222-7.
MRC Working Party on Children Conceived by In Vitro Fertilization. Births in Great Britain resulting from assisted conception 1978-87. BMJ (Clinical Research ed.) 1990;300:1229-33.
Mushin DN, Barreda-Hanson MC, Spensley JC. In vitro fertilization children: early psychosocial development. Journal of In Vitro Fertilisation and Embryo Transfer, 1986;3(4):247-52.
Petersen K, Hornnes PJ, Ellingsen S, et al. Perinatal outcome after in vitro fertilisation. Acta Obstetricia et Gynecologica Scandinavica, 1995;74:129-31.
Raoul-Duval A, Bertrand-Servais M, Lenir-Könirsch H, Frydman R. Psychological follow-up of children born after in-vitro fertilization. Human Reproduction, 1994;9:1097-1101.
Reubinoff BE, Samueloff A, Ben-Haim M, Friedler S, Schenker JG, Lewin A. Is the obstetric outcome of in vitro fertilized singleton gestations different from natural ones? A controlled study. Fertility and Sterility, 1997;67:1077-83.
Ron-El R, Lahat E, Golan A, et al. Development of children born after ovarian superovulation induced by long-acting gonadotropin-releasing hormone agonist and menotropins, and by in vitro fertilization. Journal of Pediatrics, 1994;125:734-7.
Rufat P, Olivennes F, Mouzon de J, Dehan M, Frydman R. Task force report on the outcome of pregnancies and children conceived by in vitro fertilization. Fertility and Sterility, 1994;61:324-30.
Stewart AL, Rifkin L, Amess PN et al. Brain structure and neurocognitive behavioural function in adolescents who were born very preterm. Lancet 1999;353:1653-7.
Tan SL, Doyle P, Campbell S, et al. Obstetric outcome of in vitro fertilization pregnancies compared with normally conceived pregnancies. American Journal of Obstetrics and Gynecology, 1992;167:778-84.
Van der Meulen BF, Smrkovsky M. Bayley ontwikkelingsschalen. Thesis, State University Groningen, 1984.
Wang JX, Clark AM, Kirby CA, et al. The obstetric outcome of singleton pregnancies following in-vitro fertilization/gamete intra-Fallopian transfer. Human Reproduction, 1994;9:141-46.
Wolke D, Ratschinski G, Ohrt B, Riegel K. The cognitive outcome of very preterm infants may be poorer than often reported: an empirical investigation of how methodological issues make a big difference. Eur J Pediatr, 1994;153:906-15.

Chapter 7

Couples' expectations of having a baby through IVF
Research letter

S.E. Buitendijk[1], M. Fekkes[1], D.D.M. Braat[2], G.H.W. Verrips[1]

1 TNO Prevention and Health, Division of Child Health, Leiden, PO Box 2215, 2301 CE, the Netherlands.

2 Department of Obstetrics and Gynaecology, University Hospital St. Radboud, Nijmegen, PO Box 9101, 6500 HB Nijmegen, the Netherlands.

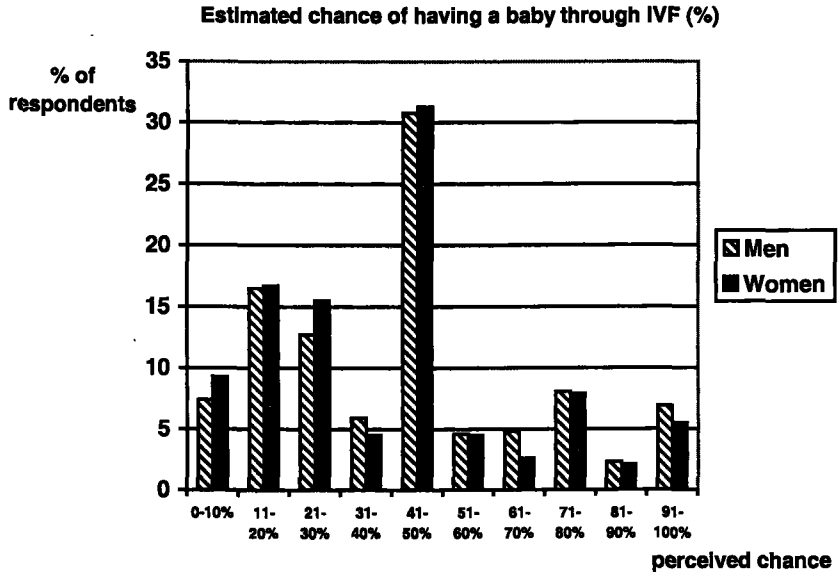
Submitted for publication

Summary

Dutch couples' perceptions of the chance of having a baby after IVF appear to be realistic. Most respondents estimated their chance to be around 50% and take into account the type of their infertility and female partner's age.

The probability of In Vitro Fertilisation (IVF) treatment resulting in a live birth is an important statistic for IVF physicians as well as for IVF couples entering into the programme. In the Netherlands, as in other countries, couples starting IVF are being counselled by their physician about the chances of the procedure's success. However, a paucity of reliable data exists to predict individual couples' success. Follow-up studies of IVF patients that calculate success per treated couple, as well as studies that suggest/show (theoretical) cumulative pregnancy rates, report pregnancy rates of around 50% after three treatment cycles (Bergh 1995, Roest 1995, Engmann 1999). Several studies show that older age of the female partner, longer duration of infertility, and male or multiple factor infertility negatively influence the chance of a viable pregnancy and subsequent live birth after IVF (Tan 1992, Kahn 1993, Panthos 1998). Information more specific to an individual couple's chance of success does not appear to be available before the start of the treatment. Prognostic models, taking into account a variety of couple and clinic factors, have shown to be of no use (Stolwijk 1998), and IVF physicians show little concordance in their estimates of chance of pregnancy and live birth after IVF for couples with different medical histories (Wiegerinck 1999).

As part of a larger study, we studied couples' perceived chances of having a baby through IVF. Six large IVF centres took part in this study. During 1997 and 1998, 493 couples who were about to enter into the treatment programme and had not previously undergone IVF were asked to participate. Each partner of the couple was asked individually to fill out a number of written questionnaires. There were 867 respondents (419 couples and 29 individual men and women). One of the questions pertained to what the respondent felt his or her percentage chance was of having a baby through IVF (duration of treatment or number of cycles anticipated was not specified). The average estimated chance of having a baby through IVF was 44.7%. Among respondents, 30% (n=224) estimated their chance to be exactly 50%, 0.7% (n=5) to be 0%, 4.3% (n=32) to be 100% and 8.1% (n=61) estimated their chance to be higher than 80%. Graph 1 shows the distribution of the perceived chance for men and women. Further, we studied differences in estimates between men and women of the same couple. In 30% of all couples, men and women estimated their chance of having a baby through IVF exactly the same. In 75% of couples, estimates did not differ more than 25 percentage points from each other. In 6% of couples, the discrepancy between the man and the woman was more than 50 percentage points.



We related respondents' estimates of success to their sex, to duration of their infertility (defined as time between first trying to become pregnant and day of filling out the questionnaire), to type of infertility (tubal only, male subfertility, idiopathic, or a combination) and to age of the female partner. We used a linear regression model, taking into account all factors simultaneously (table 1). Age of the woman and type of infertility appear to influence the estimate of success, while sex of the respondent and duration of infertility does not. For instance, with each increase in year of the woman's age, the estimated chance of success decreases with by 1.0 percentage point while presence of sperm abnormalities decreases the estimated chance of success by 10.0 percentage points as compared to diagnosed tubal pathology.

Table 1 Linear regression analysis of the effect of a number of factors on respondents estimates of the chance of having a baby through IVF (n=750)

Factor	n (%)	Difference in estimate, in % points [95% CI]	p - value
Sex of respondent			
male	362 (48.3)		
female	388 (51.7)	-2.8 [-6.4 to 0.7]	0.121
Woman's age (years)		- 1.0 [-1.4 to -0.6]	< 0.0001
Duration of infertility (years)		0.4 [-0.3 to 1.0]	0.302
Type of infertility			
tubal	101 (13.5)		
male subfertility	253 (33.7)	-10.0 [-15.8 to - 4.1]	0.001
idiopathic	163 (21.7)	- 8.7 [-14.9 to - 2.6]	0.006
combination	233 (31.1)	- 10.3 [-16.2 to - 4.5]	0.001

We conclude that, overall, Dutch couples have realistic expectations of their chances of having a baby after IVF. In the Netherlands, estimated success after three treatments is a useful outcome measure since the first three IVF treatments are paid for by health insurance. This percentage success is estimated to be around 50%, which is the most frequently reported respondents' estimate as well. Respondents are aware of the relationship between their medical history and the chance of success, with the exception of the factor 'duration of infertility'. Of partners in a couple 70% were not in concordance with respect to their chance estimate, although most couples are counselled together by their IVF physician. This indicates that respondents use additional sources of information and/or that men and women use different interpretations of the data to determine their final estimate of success. However, only 6% of couples were very discrepant (over 50 percentage points different) in their estimate.

A realistic estimate of success by individual partners as well as within couples is likely to be an important psychological factor while entering into an IVF treatment. Dutch couples appear to be well enough informed in that respect.

References

Bergh C, Josefsson B, Nilsson L, Hamberger L. The succes rate in a Swedish in-vitro fertilization unit: a cohort study. Acta Obstet Gynecol Scand 1995;74(6):446-50.

Engmann L, Maconochie N, Bekir JS, Jacobs HS, Tan SL. Cumulative probability of clinical pregnancy and live birth after a multiple cycle IVF package: a more realistic assessment of overall and age-specific success rates ? Br J Obstet Gynaecol 1999;106(2):165-70.

Kahn JA, Von During V, Dunde A, Sordal T, Molne K. The efficacy and efficiency of an in-vitro fertilization programme including cryopreservation: a cohort study. Human Reprod 1993;8(2):247-52.

Pantos K, Athanasiou V, Stefanidis K, Stavrou D, Vaxevanoglou T, Chronopoulou M. Influence of advanced age on the blastocyst development rate and pregnancy rate in assisted reproductive technology. Fertil Steril 1998;71(6) 1144-6.

Stolwijk AM, Straatman H, Zielhuis GA, Jansen CA, Braat DD, Van Dop PA, Verbeek AL. External validation of prognostic models for ongoing pregnancy after in-vitro fertilization. Hum Reprod 1998;13(12):3542-9.

Tan SL, Doyle P, Campbell S, et al. Obstetric outcome of in vitro fertilization pregnancies compared with normally conceived pregnancies. American Journal of Obstetrics and Gynecology 1992;167:778-84.

Wiegerinck MA, Bongers MY, Mol BW, Heineman MJ. How concordant are the estimated rates of natural conception and in-vitro fertilization/embryo transfer succes? Hum Reprod 1999;14(3):689-93.

Chapter 8

Health-related quality of life in relation to gender and age in couples planning IVF treatment

M. Fekkes[1], S.E. Buitendijk[1], G.H.W. Verrips[1], D.D.M. Braat[2], A.M.A. Brewaeys[3], J.G. Dolfing[4], M. Kortman[5], R.A. Leerentveld[6], N.S. Macklon[7]

1 TNO Prevention and Health, Division of Child Health, PO Box 2215, 2301 CE Leiden, the Netherlands.
2 Department of Obstetrics and Gynaecology, University Hospital St Radboud, PO Box 9101, 6500 HB Nijmegen, the Netherlands.
3 Department of Obstetrics and Gynaecology, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, the Netherlands.
4 Department of Obstetrics and Gynaecology, Diaconessen Hospital, Postbus 998, 2270 AZ Voorburg, the Netherlands.
5 Department of Obstetrics and Gynaecology, University Medical Center, PO Box 85500, 3508 GA Utrecht, the Netherlands.
6 Department of Obstetrics and Gynaecology, Isala klinieken, Location Sophia, PO Box 10400, 8000 GK Zwolle, the Netherlands.
7 Department of Obstetrics and Gynaecology, Erasmus University Medical Center, PO Box 2040, 3000 CA Rotterdam, the Netherlands.

Abstract

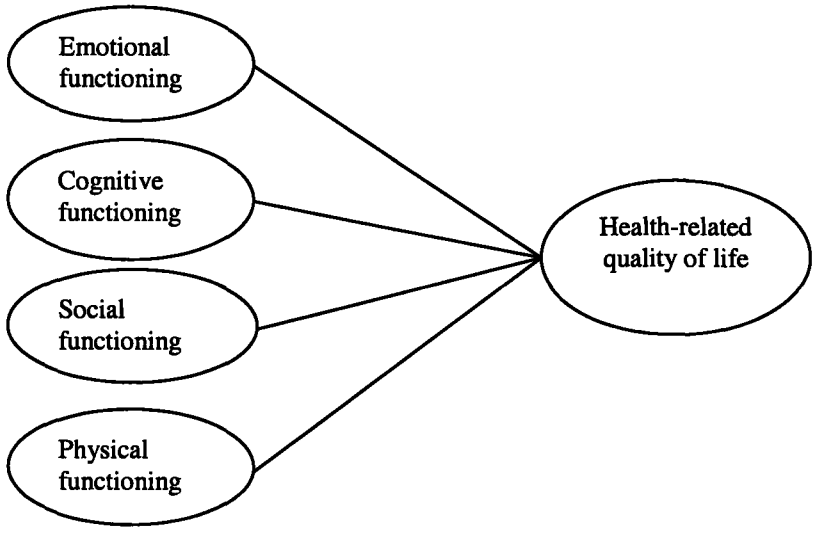
Infertility and IVF treatment have been reported to negatively influence the psychosocial status of patients. Suboptimal psychological functioning may lead to discontinuation of the IVF treatment. It is unclear whether certain subgroups may have more problems and if so, whether these problems occur in specific domains. We employed the concept of health-related quality of life to study in which specific domains people planning IVF may report more problems than the general population does. The concept of health-related quality of life includes four domains: emotional, cognitive, social and physical functioning. We compared different age subgroups as well as separate subgroups of 425 men and 447 women planning IVF treatment to comparable population subgroups.
In the domains cognitive and physical functioning no significant differences were found between any subgroup of IVF patients compared to the general population. However, the results for the domains emotional and social functioning showed that the youngest age-group of IVF women (21-30 years old) has worse scores. This subgroup of women also differed in a number of background variables from the older group of IVF women. They have lower education and report more irrational parenthood cognitions. This last observation indicates they have a stronger sense than other IVF women do that they need to have a child in order to live a happy life. For this particular group a short cognitive counseling therapy should be developed, directed towards changing irrational cognitions in order to diminish their negative emotional impact.

Introduction

Infertility and it's treatment have a considerable impact on a person's quality of life. Research suggests that infertility problems are among the most upsetting experiences in peoples' lives (Freeman 1985, Mahlstedt 1987). Most couples who plan to have IVF treatment have already experienced extensive and emotionally challenging methods of diagnosis and/or treatment. Often, their remaining treatment options are limited. The IVF procedure is considered by many couples to be the 'end of the line' (Shaw 1988). People who seek IVF treatment have been reported to be more anxious and emotionally distressed than people in the general population (Berg 1991, Newman 1991, Lukse 1999). It is still unclear whether this heightened level of distress occurs in all couples planning to undergo IVF treatment. Certain subgroups may have more problems and further, these problems may occur in specific domains. It is not unlikely, for instance, that infertility has a different impact on young people than on relatively older people planning to undergo IVF treatment. In the Netherlands, older couples cannot apply for adoption of a baby or young child. For these couples the IVF treatment may be their last possibility for having a baby. This may in turn cause excess stress during their treatment. Several studies suggest that the impact of infertility and its treatment is higher in women than in men (Freeman 1985, Callan 1988, Newton 1990, Collins 1992, Newton 1999, Oddens 1999), making it important to study men and women separately. To our knowledge the psychological status of people planning to undergo IVF has to date not been studied in different age groups and in men and women separately.
If IVF clinicians are aware at the start of the treatment which subgroups are most likely to suffer high levels of distress, they may be more able to gear their counseling efforts

towards these groups. Couples usually undergo several cycles of IVF treatment. However, if the first IVF attempt fails, a high level of emotional strain appears to be the main reason not to continue with a new IVF cycle (Mahlstedt 1987, Goverde 2000).

The concept of health-related quality of life may be a useful tool in discovering in which domains people planning IVF may have more problems than reported in the general population. Health-related quality of life as perceived by the patients themselves is increasingly used to measure the impact of disease and the effects of treatment. It is widely accepted that health-related quality of life includes the perception of patients of their functioning in four domains: physical functioning, emotional functioning, social functioning, and cognitive functioning (Aaronson 1988; Verrips 1999).



In the present study we adopted this concept of health-related quality of life. We included the four domains in which a difference may occur between the group planning to undergo an IVF treatment and the general population. To measure these domains we used validated, generic and normed instruments that were available in the Dutch language.
The main research question we aimed to answer is whether health-related quality of life in couples undergoing IVF treatment differs from that in the general population, comparing different age subgroups as well as separate subgroups of men and women.

Material and methods

Study participants

Couples with infertility problems were recruited from six IVF clinics in the Netherlands between January 1996 and June 1997. Couples were eligible to participate if they spoke Dutch, planned to undergo IVF treatment at one of the six IVF clinics and were considered suitable by the IVF clinic to undergo IVF treatment. Couples were not included in the study if they had previously undergone IVF treatment or planned to undergo ICSI-treatment.

Procedure

During their visit to the IVF clinic couples were informed by their IVF physician about the aims of the study and invited to participate. Couples that agreed to participate filled out a written informed consent form. Participants were subsequently mailed a set of questionnaires to their home. Each partner received a separate set of questionnaires. Completion of the questionnaires lasted on average around 45 minutes. Respondents could send the completed questionnaires back to the researchers free of charge. A reminding letter was mailed in case a participant did not return the questionnaires within 4 weeks. Participants who did not respond within 8 weeks were approached by telephone and asked again if they were willing to complete the questionnaires.

Measures

Four outcome measures as outlined below were used in the study. Normative data from the general population were available for all measures except for the Irrational Parenthood Cognitions-Scale, which was developed by the authors.

The Hopkins Symptom Checklist (HSCL) (Derogatis 1975).
This 57-item measure provides an assessment of psychological and physical discomfort. Apart from an overall scale, two subscales can be derived from this measure. The 17-item 'psychological complaints'-subscale aims to measure psychological and neurotic complaints such as unpleasant thoughts, outbursts of anger, worrying, and despair. The 8-item 'somatic complaints'-subscale aims to measure physical complaints such as headaches, dizziness, palpitations and backache. Normative data are derived from Luteijn (1984, 1999).

The Sickness Impact Profile (SIP) (Bergner 1981; Jacobs 1990; Jacobs 1992)
The SIP is a 136-item measure that determines the influence of illness and/or health complaints on daily functioning. Twelve subscales can be derived from this measure. These scales may be divided into two dimensions: 'physical dysfunctioning' and 'emotional-social functioning'. Only the scales of the emotional-social dimension were used in this study. These are the scales 'emotional behavior', 'social interactions', 'alertness behavior' and 'communication'. Normative data are derived from Jacobs *et al.* (1990).

The Irrational Beliefs Inventory (IBI) (Koopmans 1994).
This 50-item measure assesses irrational cognitions and is based on the Irrational Beliefs Test (IBT) (Jones 1969) and the Rational Behavior Inventory (RBI) (Shorkey 1977). This measure can be of use in cognitive interventions. Within the total of 50 IBI-items five subscales exist which assess 'worrying' ('I worry about certain future things'), 'rigidity' ('There is only one right way to do things'), 'need for esteem' ('I want everybody to like me'), 'external control' ('If a person wants to, he can be happy under almost any circumstance'), and 'problem avoidance' (I avoid facing my problems'). Normative data are derived from the manual (Timmerman 1993).

The Irrational Parenthood Cognitions (IPC-)Scale.
A questionnaire was developed by the authors to measure specific irrational cognitions concerning the need to have children in order to live a happy life. This scale contains 14 items. Examples are: "a life without children is useless and empty" or "you start hating your body when you cannot have children". Respondents were asked to score on a 5-point-scale to what extent they agreed with these statements. The items were subsequently totaled to a 0-56 scale-score, with higher scores indicating a stronger need to have children in order to live a happy life.
The reliability of this scale, measured by Cronbach's alpha, was 0.84 for male respondents and 0.87 for female respondents. An English version of the IPC-scale is presented in Appendix A.

Background variables.
With a separate questionnaire several background variables were measured such as: infertility diagnosis, duration of infertility, history of infertility treatments, presence of children in the family at the start of the IVF treatment and education and age of both partners.

The following (sub)scales were used to measure the different dimensions of the concept of health-related quality of life.
Emotional functioning: The SIP-subscale 'emotional behavior' and the HSCL-subscale 'psychological complaints'.
Cognitive functioning: The IBI-total scale.
Social functioning: The SIP-subscales 'social interactions', 'alertness behavior', and 'communication'.
Physical functioning: The HSCL-subscale 'somatic complaints'.

Analyses

Normative data from the general population were only available to us in tables with means and standard error or standard deviation. This limited the options for analysis.
Two-sided Student's *t*-test analyses compared the mean scale scores on the outcome variables for the IVF sample with the scores for the general population. This analysis was carried out separately for men and women. The analyses were subsequently carried out for different age groups (21-30, 31-40, and 41-50 years).

The group of men over 50 years of age and the group of women over 40 years of age were not included in the analysis since they were too small to allow a meaningful analysis to be performed. Due to missing data the number of respondents included in the analyses differed per outcome measure.

Results

Characteristics of the sample
A total of 524 couples agreed to participate in the study by filling out the informed consent form. Of this group, 425 men and 447 women (420 couples plus 32 individual men and women) returned the questionnaire. This yields an 82% response rate (84% of the women and 80% of the men). The mean age for women was 32 years (range 21-43) and for men 36 years (range 22-59 years).

Non-response analysis
The non-responders did not differ from the responders with regards to age, duration of infertility and the presence of children in the family. A difference did exist between the two groups in the distribution of the infertility diagnosis. The non-response group included more women with tubal pathology (25% vs. 14%) and fewer patients with idiopathic infertility (10% vs. 21%).

Characteristics of the sample
In table 1 and table 2 the distribution of background variables is presented for different age groups of men and women separately.

Table 1 Mean scores and ANOVA p-values for duration of infertility and Irrational Parenthood Cognitions-scale. A higher score on the Irrational Parenthood Cognitions-scale indicates a stronger sense that becoming a parent is the major focus in life.

	Women			Men			
	21-30 n = 144	31-40 n = 289	p-value	21-30 n = 70	31-40 n = 251	41-50 n = 65	p-value
Duration of infertility (yr.)	4.0	4.6	.036	3.8	4.4	4.8	.053[a]
Irrational Parenthood Cognitions (IPC)-score	34.1	29.2	<.001	29.2	25.8	24.3	.024[b]

[a] Posthoc analysis Tuckey HSD significance level between age groups 21-30 and 31-40: p = .164; between groups 31-40 and 41-50: p = .456 ; and between groups 21-30 and 41-50: p = .044

[b] Posthoc analysis Tuckey HSD significance level between age groups 21-30 and 31-40: p = .060; between groups 31-40 and 41-50: p = .557 ; and between groups 21-30 and 41-50: p = .024

Table 2 Percent distribution and p-value of Chi-square for the variables education, children present in the family before the start of the IVF treatment, history of other infertility treatments in the past and infertility diagnosis, for different age groups of IVF men and -women

	Women			Men			
	21-30 n = 144	31-40 n = 289	p-value	21-30 n = 70	31-40 n = 275	41-50 n = 72	p-value
Education							
- low	32%	26%	.001	39%	35%	33%	.212
- medium	56%	45%		45%	35%	33%	
- high	12%	29%		16%	30%	33%	
Children present in family							
- no	91%	76%	<.001	89%	82%	62%	<.001
- yes	9%	24%		11%	18%	38%	
Other infertility treatment before IVF treatment.							
- no	39%	34%	.327	36%	37%	28%	.464
- yes	61%	66%		64%	63%	72%	
Infertility diagnosis							
- tubal pathology	6%	15%	.001	13%	13%	19%	.503
- sperm abnormalities	44%	31%		41%	32%	36%	
- idiopathic infertility	14%	25%		16%	23%	17%	
- combination	36%	29%		30%	32%	28%	

A number of differences in background variables existed between the different age groups of men and women, as is shown in table 1 and table 2. For both IVF- men and -women, the youngest age group (21-30 years old) had a shorter duration of infertility and indicated more Irrational Parenthood Cognitions. Participants in this age group less often had children at the start of the first IVF treatment. Women in this youngest age group more often had a lower education.

Emotional functioning

Emotional functioning was assessed with the 'emotional behavior'-scale of the SIP, and with the 'psychological complaints'-subscale from the HSCL. Mean scores on these subscales are presented in table 3 and table 4, for both the IVF population and the general population.

Table 3 Mean scores, standard error (SE), T-value and p-value, for the Sickness Impact Profile (SIP)-scale 'Emotional behavior' for both the IVF population and the general population. Higher scores indicate worse functioning.

	IVF population		General population			
	N	Mean (SE)	N	Mean (SE)	T-value	p-value
Emotional Behavior						
- Women 21 – 30	145	15.3 (1.4)	106	4.0 (0.7)	7.17	< 0.001
- Women 31 – 40	291	11.1 (0.9)	53	3.2 (1.2)	3.44	< 0.001
- Men 21 – 30	70	4.8 (1.3)	86	1.7 (0.6)	2.18	< 0.05
- Men 31 – 40	274	3.6 (0.5)	54	7.3 (2.0)	-1.80	ns
- Men 41 – 50	79	6.5 (1.4)	34	4.9 (2.3)	0.60	ns

Table 4 Mean scores, standard error (SE), T-value and p-value, for the Hopkins Symptom Checklist (HSCL)-scale 'Psychological complaints' for both the IVF population and the general population. Higher scores indicate worse functioning.

	IVF population		General population			
	N	Mean (SE)	N	Mean (SE)	T-value	p-value
Psychological complaints						
- Women 21 – 30	143	9.2 (0.6)	58	10.5 (1.1)	-1.09	ns
- Women 31 – 40	288	7.1 (0.4)	64	8.1 (1.0)	-1.04	ns
- Men 21 – 30	68	4.9 (0.7)	36	5.6 (0.8)	-0.61	ns
- Men 31 – 40	267	4.5 (0.3)	70	5.0 (0.6)	-0.92	ns
- Men 41 – 50	71	5.6 (0.8)	68	5.7 (0.8)	-0.03	ns

The SIP-scores for the subscale 'Emotional behavior' in table 3 show that IVF women exhibited more emotional behavior problems than women in the general population. Within the youngest age group this difference was the largest. For the IVF men, the youngest age group demonstrated significantly more emotional behavior problems compared with the general population. However, the older age groups of IVF men did not differ significantly from the general population. The HSCL scores for the subscale 'Psychological complaints in table 4 show that the IVF women did not have more psychological complaints than the women from the general population. This was the case for both age groups.

Physical functioning

Perceived physical functioning was assessed with the 'psychosomatic complaints' scale from the HSCL. Mean scores on this scale are being presented in table 5, for both the IVF population and the general population.

Table 5 Mean scores, standard error (SE), T-value and p-value, for the Hopkins Symptom Checklist (HSCL)-scale 'Psychosomatic complaints' for both the IVF population and the general population. Higher scores indicate less good functioning.

	IVF population		General population			
	N	Mean (SE)	N	Mean (SE)	T-value	p-value
Psychosomatic complaints						
- Women 21 – 30	144	3.5 (0.2)	58	4.0 (0.5)	-1.01	ns
- Women 31 – 40	289	2.8 (0.2)	64	3.6 (0.4)	-1.93	ns
- Men 21 – 30	69	1.4 (0.2)	36	2.7 (0.5)	-2.09	< 0.01
- Men 31 – 40	267	1.5 (0.1)	70	1.9 (0.3)	-1.69	ns.
- Men 41 – 50	71	2.1 (0.3)	68	2.5 (0.4)	-0.78	ns

The HSCL scores for the subscale 'Psychosomatic complaints' in table 5 show that the IVF women did not differ from the general population. This was the case for both age groups. In IVF men the age group '21-30 years' reported fewer psychosomatic complaints than the general population. For the other two age groups no significant differences were found.

Cognitive functioning

Cognitive functioning was measured with the IBI-total scale. Mean scores on this scale are being presented in table 6.

Table 6 Mean scores, standard deviation ± SD, T-value and p-value, for the Irrational Beliefs Inventory (IBI)-scale for both the IVF population and the general population. Higher scores indicate less optimal functioning.

	IVF population		General population			
	N	Mean ± SD	N	Mean ± SD	T-value	p
IBI-total						
- Women 21 – 30	142	143.0 ± 17.9	72	144.0 ± 13.4	-0.46	ns
- Women 31 – 40	284	136.8 ± 18.0	47	137.0 ± 17.4	-0.07	ns
- Men 21 – 30	69	133.5 ± 12.7	38	131.6 ± 17.5	0.64	ns
- Men 31 – 40	265	131.1 ± 14.9	37	135.0 ± 17.4	-1.29	ns
- Men 41 – 50	70	132.6 ±20.6	42	139.4 ± 16.4	-1.92	ns

The IBI-scores in table 6 show that no significant differences existed between the IVF group and the general population.

Social functioning

Social functioning was measured with three SIP-subscales, i.e. social interaction, alertness behavior and communication. The mean scores on these scales for both the IVF- and the general population (Jacobs 1990) are presented in table 7.

Table 7 Mean scores, standard error (SE), T-value and p-value, for the Sickness Impact Profile (SIP)-scales 'Social interaction', 'Alertness behavior' and 'Communication' for both the IVF population and the general population. Higher scores indicate worse functioning.

	IVF population		General population		T-value	p-value
	N	Mean (SE)	N	Mean (SE)		
Social interaction						
- Women 21 – 30	144	11.0 (0.9)	106	2.5 (0.5)	8.01	< 0.001
- Women 31 – 40	289	7.9 (0.7)	53	6.2 (1.5)	1.04	ns
- Men 21 – 30	70	6.3 (1.3)	86	1.7 (0.4)	3.32	< 0.001
- Men 31 – 40	274	5.8 (0.5)	54	6.1 (1.4)	-0.20	ns
- Men 41 – 50	72	8.0 (1.5)	34	5.1 (1.8)	1.37	ns
Alertness behavior						
- Women 21 – 30	144	13.5 (1.3)	106	3.6 (0.8)	6.47	< 0.001
- Women 31 – 40	289	11.9 (1.0)	53	5.8 (1.8)	2.96	< 0.01
- Men 21 – 30	70	6.4 (1.6)	86	1.5 (0.6)	2.94	< 0.01
- Men 31 – 40	274	6.0 (0.8)	54	8.9 (2.6)	-1.08	ns
- Men 41 – 50	72	7.9 (1.9)	34	4.5 (1.9)	1.53	ns
Communication						
- Women 21 – 30	144	3.7 (0.8)	106	0.5 (0.3)	3.82	< 0.001
- Women 31 – 40	289	3.1 (0.5)	53	0.5 (0.4)	4.32	< 0.001
- Men 21 – 30	70	3.6 (1.3)	86	0.8 (0.3)	2.11	< 0.05
- Men 31 – 40	274	3.1 (0.6)	54	3.3 (1.2)	-0.15	ns
- Men 41 – 50	72	6.3 (2.0)	34	0.6 (0.4)	3.18	< 0.01

The SIP-scores in table 7 show that IVF women had more problems in social functioning than women in the general population. For younger women especially, large differences were shown on all the three subscales "social interaction", "alertness", and "communication". Older IVF women reported more problems for two of these subscales, namely "alertness" and "communication".
The youngest group of IVF men differed from the general population in a similar way to the youngest group of IVF women. On all three social functioning subscales young IVF men indicated higher problem scores than the general population. The older IVF men (31-40 and 41-50 years) did not differ on most of the social functioning subscales. Only on the subscale "communication" did the oldest group of IVF men show more problems than the general population.

Correlation between the scales

In table 8 the correlation coefficients between the outcome measures are presented.
The correlation coefficients in table 8 show that the relationships between the outcome measures are fair to moderate. Most correlation coefficients vary between 0.30 and 0.60. Correlations between the Irrational Parenthood Cognitions and emotional and psychological complaints are moderate.

Table 8 Pearson Correlation Coefficients between the scales. Coefficients for men (n = 423) are presented under the diagonal, coefficients for women (n = 444) are presented above the diagonal.

	SIP-e	HSCL-pc	HSCL-sc	IBI	IPC	SIP-si	SIP-ab	SIP-c
SIP-Emotional (SIP-e)	-	.64	.42	.48	.46	.63	.51	.42
HSCL-psychological complaints (HSCL-pc)	.59	-	.51	.56	.49	.63	.63	.43
HSCL-somatic complaints (HSCL-sc)	.50	.53	-	.30	.34	.38	.37	.30
Irrational Cognitions-total (IBI)	.39	.49	.30	-	.54	.41	.38	.32
Irrational Parenthood Cognitions (IPC)	.34	.34	.15	.50	-	.43	.36	.35
SIP-Social interaction (SIP-si)	.61	.57	.34	.33	.36	-	.59	.41
SIP-Alertness behavior (SIP-ab)	.42	.56	.35	.35	.24	.50	-	.47
SIP-Communication (SIP-c)	.47	.50	.38	.33	.27	.35	.49	-

Man-woman correlation within IVF couples

In table 9 the correlation coefficients between men and women within the couples seeking IVF treatment are being presented.

Table 9 Pearson Correlation coefficients (r) and explained variance, within IVF couples.

	N (pairs)	R	Explained variance
Emotional functioning:			
- emotional behavior (SIP)	419	.29	8 %
- psychological complaints (HSCL)	404	.27	7 %
Physical functioning:			
- somatic complaints (HSCL)	405	.14	2%
Cognitive functioning:			
- irrational cognitions (IBI-total)	406	.31	10%
- irrational parenthood cognitions (IPC)	365	.55	30%
Social functioning:			
- social interaction (SIP)	419	.44	19 %
- alertness behavior (SIP)	419	.24	6 %
- communication (SIP)	419	.35	12 %

The correlation coefficients in table 9 show no strong relationship between the scores of men and women within a couple. Most coefficients have a score below 0.35, which corresponds with an explained variance of below 12%. This indicates that for most outcome measures less than 12% of a woman's scale-score can be explained by the score of her male partner. For the irrational parenthood cognitions and the SIP-social interaction scale, the correlation and the explained variance were substantially higher.

Discussion

We investigated health-related quality of life of a large group of men and women planning to undergo IVF treatment. Four domains of health-related quality of life were measured, namely perceived emotional, physical, cognitive, and social functioning. The scores of the IVF group were compared to normative data from the general population. We found that IVF women, in particular young women aged 21 to 30, had more social and emotional problems than women of the same age group in the general population. The group of young IVF men (aged 21 to 30) also, reported more social and emotional problems than a similar age group of men in the general population. These differences were somewhat smaller for men than for women. We found no difference in social and emotional functioning between the groups of relatively older IVF men and the general population. No substantial differences were found in cognitive and physical functioning for all age groups of IVF- men and women compared to the general population.

The conclusion that especially younger IVF-men and women have more problems may at first sight appear somewhat surprising. At the onset of our study we expected that older couples may have more stress-related problems since for most of these couples IVF offers their very last chance of having a baby. Younger couples, on the other hand, have more time and may still have the option to apply for adoption. However, our results show that younger couples particularly, indicated more stress-related problems. One explanation may be that these couples respond differently to the process of infertility diagnostic tests and/or treatments that they went through before entering into the IVF program. It is possible also that the infertility itself is more stressful to them than to older couples. In order to be in an IVF program before the age of 29 a woman must have started trying to become pregnant relatively early in life, especially by Dutch standards. Dutch women have the highest age at the birth of their first child of all women in Western countries, with an average age of 29.0 in 1997 when this study was performed (CBS 1999). Most women had tried to become pregnant for several years before entering into an IVF program. Therefore, the group of young IVF women had a first attempt at starting a family at an earlier age than most Dutch women. Perhaps their outlook on a life without children is fundamentally different from that in older IVF couples. If indeed for these young IVF women raising children is their single most important role in life, not being able to have children may be more stressful to them than to older women who may have had other priorities in the period of their lives before trying to start a family. The results of our study indicate that the younger IVF patients are a different group when compared with the older patients. Less often they already had children present in their family. The women in this younger age group had a lower educational level than IVF women in the older age groups and both men and women in this group showed more irrational parenthood cognitions than older IVF couples did. This last result indicates that these patients have a stronger notion than other IVF couples that they need to have children in order to live a happy life. This intense focus on having a child is found in other studies to be the predominant factor in anticipated stress of IVF treatment for both males and females (Collins 1992). Newton (1999) found that both men and women with fertility problems who have a high need for parenthood and a strong rejection of a childfree lifestyle also have more symptoms of depression and anxiety.

It was shown from our data that within the subgroup of younger IVF couples especially the women reported lower health-related quality of life. This finding has been confirmed by other studies which indicate that for women, much more often than for men, infertility is one of the most upsetting experiences of their lives (Freeman 1985) and is related to higher levels of anxiety (Slade 1997). However, the results of our study also show that IVF patients, both men and women, do not have higher levels of cognitive or (psycho)somatic problems, nor higher levels of psychological complaints on the HSCL-scale. In addition, relatively older IVF men did not differ from the general population with regard to emotional and social functioning. It can be concluded that most subgroups of IVF patients do not have a substantially lower health-related quality of life than the general population.

The higher levels of problematic functioning for IVF women when compared with the general population are found only on the emotional and social functioning scales of the SIP. This questionnaire determines the influence of illness and/or health complaints on daily functioning, indicating that this problematic functioning may be a 'short term' effect of the treatment that these patients are anticipating. Even if this is the case these higher levels of problematic functioning should be taken seriously. IVF treatments during which people feel anxiety and depression may lead to avoidance coping. When a first IVF attempt fails, negative feelings and avoidance may lead to discontinuing the treatment (Cook 1989). Mahlstedt (1989) found that the main reason for couples not continuing IVF treatment was the emotional strain of the previous treatment cycle. Goverde (2000) recently reported a randomized clinical trial comparing IVF to Intra Uterine Insemination (IUI). Couples in the IVF program were less likely to achieve pregnancy although the success rate per cycle was higher for IVF than for IUI. The main reason for this difference was that IVF couples discontinued treatment before the maximum number of cycles had been completed, far more often than did couples in the other treatment arm. Counseling directed towards lowering the emotional strain may be an important measure to prevent couples from discontinuing the treatment after a failed attempt. It is advised to initiate therapeutic counseling before infertility treatment in order to provide patients with opportunities to learn the skills to cope with the infertility and the associated medical procedures (Lukse 1999). Our study showed that especially younger women starting IVF treatment may benefit from counseling for their higher levels of emotional and social problems. Since a substantial relationship exists between the irrational parenthood cognition score of men and women within a couple and since young men also reported a lower health-related quality of life on some domains, it may be advisable to offer counseling therapy to the couple. Counseling may be directed towards changing the cognitions of those patients who are focused on the idea that having a baby is necessary to live a happy life and have high scores on related notions on the IPC scale. The correlation coefficients between the irrational (parenthood) cognitions and emotional functioning especially, were high. This suggests that changing the irrational cognitions may have a positive effect on emotional functioning. Such counseling may follow the principles of Rational Emotive Therapy (Engels 1993, Ellis 1997) in which patients learn to change their cognitions in order to change the negative emotional impact of these cognitions. Irrational cognitions have in other areas shown to be highly amenable to cognitive therapy.

Longitudinal research is needed in order to determine whether counseling interventions will indeed enhance IVF couples' quality of life and will decrease the chance of discontinuation.

References

Aaronson NK. Quality of life: What is it? How should it be measured? Oncology 1988;2:69-74.
Berg BJ and Wilson JF. Psychological functioning across stages of treatment for infertility. J. Behav. Med. 1991;14(1):11-26.
Bergner M , Bobbitt R A, Carter WB, Gilson BS. The Sickness Impact Profile: Development and final revision of a health status measure. Med. Care 1981;19:787.
Callan VJ and Hennessey JF. Emotional aspects and support in in vitro fertilization and embryo transfer programs. J. Vitro Fert. Embryo Transfer 1998;5:290-5.
CBS (1999). Statistisch jaarboek 1999. Centraal Bureau voor de Statistiek. Voorburg/Heerlen, 1999.
Collins A, Freeman EW, Boxer AS, Tureck R. Perceptions of infertility and treatment stress in females as compared with males entering in vitro fertilization treatment. Fertil. Steril. 1992;57(2):350-6.
Cook R, Parsons J, Mason B, Golombok S. Emotional, marital and sexual functioning in patients embarking upon IVF and AID treatment for infertility. J. Reprod. Infant Psychol. 1989;7:87-93.
Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L. The Hopkins Symptom Checklist, a self-report symptom inventory. Behav. Science 1974;19:1-15.
Ellis A Albert Ellis on rational emotive behavior therapy. Am. J. Psychother 1997;51(3):309-16.
Engels GI, Garnefski N, Diekstra RF. Efficacy of rational-emotive therapy: a quantitative analysis. J. Consult. Clin. Psychol. 1993;61(6):1083-90.
Freeman EW, Boxer AS, Rickels K, Tureck R, Mastroianni L. Psychological evaluation and support in a program of in vitro fertilization and embryo transfer. Fertil. Steril 1985;43(1):48-53.
Goverde AJ, McDonnell J, Vermeiden JP, Schats R, Rutten FF, Schoemaker J. Intra-uterine insemination or in-vitro fertilisation in idiopathic subfertility and male subfertility: a randomised trial and cost-effectiveness analysis. Lancet 2000;355(9197):13-8.
Jacobs HM, Luttik A and Touw-Otten FWMM. Overzichtstabellen functionale status gegevens van een open populatie. Rijksuniversiteit Utrecht, Utrecht 1990.
Jacobs HM, Luttik A, Touw-Otten FWMM, Kastein M, de Melker RA. Measuring impact of sickness in patients with nonspecific abdominal complaints in a Dutch family practice setting. Med. Care 1992;30(3):244-251.
Jones R. A factored measure of Ellis' irrational beliefs system with personality maladjustment correlates. (Doctoral dissertation, Texas Technical College) Dissertation abstracts international, 1968;29(II-b),4379-80. (University Microfilms No. 69-6443, 1969).
Koopmans PC, Sanderman R, Timmerman I, Emmelkamp PMG. The Irrational Beliefs Inventory: Development and psychometric evaluation. Eur. J. Psychol. Ass. 1994;10:15-27.
Lukse MP, Vacc NC. Grief, depression and coping in women undergoing infertility treatment. Obst Gyn 1999;93(2):245-51.
Luteijn F, Hamel LF, Bouwman TK, Kok AR. HSCL Hopkins Symptom Checklist. Handleiding. Swets & Zeitlinger, Lisse, 1984.
Luteijn F.. (1999). Written communication.
Mahlstedt PP, Macduff S and Bernstein J Emotional factors and the in vitro fertilization and embryo transfer process. J.Vitro Fert. Embryo Transfer 1987;4(4):232-6.
Newman NE and Zopuves CG. Emotional experiences of in Vitro Fertilization Participants. J. Vitro Fert. Embryo Transfer 1991;8(6):322-8.

Newton CR, Hearn MT and Yuzpe AA. Psychological assessment and follow-up after in vitro fertilization: assessing the impact of failure. Fertil. Steril 1990;54:879-886.
Newton CR, Sherrard W and Glavac I. The fertility problem inventory: measuring perceived infertility-related stress. Fertil. Steril. 1999;72(1)54-6.
Oddens BJ, Tonkelaar den I and Nieuwenhuyse H. Psychosocial experiences in women facing fertility problems – a comparative survey. Hum. Reprod. 1999;14(1)255-261.
Shaw P, Johnston M and Shaw R. Counselling needs, emotional and relationship problems in couples awaiting IVF. J. Psychosom. Obst. Gynaecol. 1988;9:171-180.
Shorkey CT and Whiteman VL. Development of the Rational Behavior Inventory: initial validity and reliability. Educat. Psychol. Measurement 1977;37:527-534.
Slade P, Emery J and Lieberman BA. A prospective, longitudinal study of emotions and relationships in in vitro fertilization treatment. Hum. Reprod 1997;12(1):183-190.
Timmerman I, Sanderman R, Koopmans PC, Emmelkamp PMG. Het meten van irrationele cognities met de Irrational Beliefs Inventory: een handleiding. Noordelijk Centrum voor Gezondheidsvraagstukken, Groningen, 1993.
Verrips EGH, Vogels TGC, Koopman,HM et al. Measuring health-related quality of life in a child population. Eur. J. Public Health 1999;9:188-193.

Appendix 1 Irrational Parenthood Cognition-scale

	I agree totally				I disagree totally
1) Having a child is the most important thing in life.	1	2	3	4	5
2) A life without children is useless and empty.	1	2	3	4	5
3) It is absurd that some people can have children quite easily, while others never do.	1	2	3	4	5
4) The whole world revolves around children.	1	2	3	4	5
5) Friends have no idea what people go through who cannot have children.	1	2	3	4	5
6) It is impossible to understand that some women decide to have an abortion.	1	2	3	4	5
7) You start feeling inferior when you cannot have children.	1	2	3	4	5
8) You start hating your body when you cannot have children.	1	2	3	4	5
9) An IVF treatment is extremely heavy and painful.	1	2	3	4	5
10) Your whole world is destroyed when you/your partner have/has your/her period after replacement of the embryo's.	1	2	3	4	5
11) The waiting during an IVF cycle puts you through hell.	1	2	3	4	5
12) One's whole world is destroyed when the last IVF treatment fails.	1	2	3	4	5
13) Not having children causes lifelong suffering.	1	2	3	4	5
14) One would want to do anything to get pregnant.	1	2	3	4	5

Chapter 9

Summary and reflective conclusions

Introduction

In Vitro Fertilisation takes a prominent place in many overviews of the last century's important medical developments. Together with the introduction of antibiotics, of medical diagnostics such as ultrasound and magnetic resonance imaging, of immunisation programs, of blood transfusions and of the mapping of the human genome, it ranks among the top ten landmark discoveries that have changed the face of medicine. And rightfully so. Until IVF was introduced fertility specialists had few treatments available to them. Ovulation induction, tubal surgery and artificial insemination offered possibilities for couples who had trouble conceiving, but only for a limited proportion of them. Since IVF and ICSI have become available on a large scale, virtually every couple with a fertility problem can be offered hope of having their own biological child. In terms of it's social, ethical, legal and medical impact, IVF may arguably be the most far reaching medical development of the 20th century perhaps sharing it's top position with the human genome project.

Yet, IVF like many of today's medical technologies was developed as a new treatment with very little attention paid to social, ethical and legal impact. The founding fathers of IVF, Edwards and Steptoe, describe in their book 'A matter of life' in amazing detail what they went through both in their personal and in their professional lives to develop the technique and, finally, enable the birth of Louise Brown (Edwards and Steptoe 1980). They pay relatively little detail, however, to the possible societal consequences of their invention. Edwards describes the nights he spends in his laboratory, attempting to fertilise with his own sperm human eggs made available to him by friendly gynaecologists. The eggs were sent to him in human ovarian tissue left over from surgical procedures. He describes his struggles to develop a culture medium in which sperm can be made ready to fertilise eggs. In an attempt to achieve this, he developed a small semi-permeable box that could contain sperm cells. This box was subsequently inserted into the uteruses of patients waiting in hospital to undergo a gynaecological surgical procedure. He describes his many attempts to create a culture medium to develop the embryos. Learning of Steptoes advances with laparoscopic surgical techniques was a huge milestone in the endeavour. And together, Edwards and Steptoe describe how, after many years, their efforts are being rewarded by the successful pregnancy of Mrs. Brown, and the subsequent birth of Louise in 1978. The authors do not appear to be overly worried about the potential side effects of the procedure while working on developing and perfecting it. Edwards and Steptoe state that after Louise was born 4 paediatricians performed a total of 41 tests on her to subsequently conclude that she was a perfectly normal baby 'which we already knew'. Legal, ethical and social concerns related to the introduction of the procedure are mentioned in their book, but mainly as nuisances, hampering the efforts to disseminate the technique. Yet, the development of IVF has had far reaching consequences for reproduction. It has changed the social and medical meaning of parenthood and has created unrivalled complications for policy makers.

During the two decades since it's introduction much debate has ensued about the medicalisation of pregnancy and procreation caused by IVF. It was created by Edwards and Steptoe to treat a narrow set of indications, mainly related to occlusion or absence of

the fallopian tubes. It's initial success soon resulted in substantial broadening of the range of uses. With the expansion of it's use came a rapidly increasing need for IVF and with the subsequent wider use, a host of accompanying legal and social problems: the problem of surplus embryo's and their storage and ownership; the problem of donor oocytes and defining biological boundaries on reproductive age; the problem of surrogate motherhood; the problem of whether or not to permit the use of embryo's for research. The law in most countries has proven to be ill-equipped to deal with most of these issues, which is illustrated by the highly variable regulations internationally. Also, new developments contingent upon IVF have often occurred so rapidly, it appears almost inevitable that legal regulations fall several steps behind.

The emphasis on complex legal and ethical issues coupled with the medical focus on improving the effectiveness of the procedure have resulted in relatively little attention in the early years of the dissemination of IVF to the course and outcome of IVF pregnancies. Each of the different phases of the IVF treatment may potentially harm the developing embryo. Most of the hormonal preparations used in IVF treatments have not been studied extensively for their possible teratogenic effects. Second, it is yet unclear whether or not oocyte retrieval may cause mechanical damage to the egg. Third, fertilised eggs are being transferred into a culture medium for further growth. Substances in the culture medium may have adverse effects on the embryo. Fourth, the transfer procedure into the uterus of the embryo's once developed may cause damage. Last, the freezing of surplus embryo's may cause mechanical damage. Possible negative effects occurring in one or more of these phases may cause less than optimal development of the embryo. Further, the embryo enters into the uterine cavity in a non-natural way and possibly at a non-optimal stage in it's growth, while the endometrium has developed under abnormal, artificially stimulated circumstances. Each of these aspects may in itself cause abnormal development of the embryo which in turn may lead to perinatal problems such as growth restriction, preterm birth or congenital anomalies.

Also, in the early years the psychological status of couples undergoing the IVF procedure was not a topic of much research. It appeared to be assumed that involuntary childlessness necessarily created a much larger psychological burden than the treatment to overcome childlessness ever would.

Chapter 2: Overview of the literature on pregnancy outcomes and child development

Chapter 2 provides an overview of the research data concerning the health of the children born after IVF as published by early 1999. By the late 1990's a number of studies had appeared describing the health of children born after IVF. One of the main problems with IVF is the high rate of multiple pregnancies. Of all IVF pregnancies 25% to 30% are multiple pregnancies, of which 3% are triplet and higher order pregnancies. Thus, almost half the children born after IVF are part of a multiple pregnancy. Multiple pregnancies inherently carry a higher risk of complications of which prematurity, intrauterine growth restriction and perinatal mortality are the most important. Singleton pregnancies after IVF also appear at higher risk of adverse outcomes. The combined literature shows that

singleton IVF pregnancies are approximately two times more likely to end in preterm birth and low birth weight than normally conceived pregnancies. From these studies it cannot be concluded whether this excess risk is related to the IVF procedure per se or to factors related to the mothers' profiles. Women who give birth after IVF are on average older and more often primiparous than other mothers are. They may be of higher socio-economic status and have different risk taking behaviours (such as smoking and drinking) during pregnancy. They may suffer more from certain chronic conditions that may influence fertility as well as the outcome of pregnancy. The infertility itself, instead of the IVF procedure may cause the adverse outcomes. None of the studies described controlled for a large number of covariates at the same time, thus rendering it impossible to draw conclusions as to the cause of the higher risk. However, in terms of total risk of adverse outcome, the multiple pregnancy rate appears to be the most important factor in IVF pregnancies.

A number of studies were carried out into birth defects after IVF. None of these show an increase in IVF children. However, to study a possible increase in specific, well-defined defects larger numbers are needed or pooled analyses should be carried out.

A small number of studies have looked into the possible effects of cryopreservation of embryos. Cryopreservation and thawing involve major cellular changes and it is not clear whether these may have adverse effects on the offspring. The studies carried out so far show no adverse outcomes, although the numbers studied are too small to offer much reassurance.

Studies into child development after IVF are limited in number. Further, they include relatively small groups of children and have short periods of follow-up. Many disorders, such as school performance problems or attention deficit disorders can only be diagnosed at older ages. From this literature, no adverse effects on the development of children of IVF per se or of the prematurity related to it are apparent. A few studies show that the quality of parenting after IVF may be better than in the control group. If adverse effects of IVF on child development exist, these may be counterbalanced by the wantedness of the pregnancy and the superiority of the parenting. Thus, from the literature as available early 1999 it cannot be judged whether IVF children develop normally or whether the numbers studied have been too small to detect a difference.

Aims of the study

In 1992 an effort was launched by Dutch IVF clinicians and researchers to set up a study large enough and with a proper control group, to enable conclusions about the health of IVF babies. Children born from IVF pregnancies were to be studied directly after birth and compared to a control group. A subgroup of these IVF children would be studied at the age of two and compared to a control group in terms of their psychomotor development. In a separate group of study subjects, the effects of the procedure on psychological status and quality of life were to be studied. The research questions were as follows:

- do IVF children have birth weights and gestational ages that differ from naturally conceived children?
- do IVF children born after cryopreservation have different outcomes than those born after IVF without cryopreservation?
- is psychomotor development of two year old IVF children different from that of normally conceived children?
- what is the impact of IVF on quality of life in couples undergoing the treatment?

Results of the study

Chapter 3: Preterm birth

In 2043 IVF pregnancies from which 2636 children were born, the preterm birth rate was studied. It was compared to the preterm birth rate in two separate control groups, the SMOCC- (Social Medical Survey of Children attending Child Health Clinics) and the LVR- (National Obstetric Database) control group. IVF mothers were of lower parity, of higher education and less likely to belong to an ethnic minority than women in the two control groups. They were less likely to have had pre-pregnancy diabetes or hypertension than women in the SMOCC-cohort. They smoked less and consumed less alcohol than women in the SMOCC-cohort. In the IVF group the rates of induced labour and elective caesarean section were higher than those in the LVR-control group. Women delivered less often at home. Of the IVF pregnancies 26.5% were multiple, compared to 1.6% in both control cohorts. Preterm delivery occurred more often in the IVF cohort, 23.8% vs. 6% in the SMOCC and 7.6% in the LVR-cohort respectively. Singleton IVF pregnancies ended preterm more frequently than control pregnancies did (9.3% vs. 5.2% and 6.9% in the SMOCC and LVR-cohorts respectively). Singleton IVF children had an increased risk of perinatal death (3.0% vs. 1.5% in the LVR-cohort). Logistic regression analyses were carried out to correct for a large number of variables simultaneously. After correcting for multiplicity, age, parity, ethnicity, mother's education, smoking or alcohol use and pre-pregnancy diabetes and hypertension, in IVF exposed women the odds ratio of the pregnancy ending before 37 weeks was 1.6 (95% CI [1.2 – 2.1]) compared to SMOCC-women. After correcting for multiplicity, age, parity, ethnicity, pre-pregnancy diabetes and hypertension, and whether or not labour was induced, in IVF exposed women the odds ratio of the pregnancy ending before 37 weeks was 1.2 [95% (CI 1.1 – 1.4)] compared to LVR-women. The OR for preterm birth before 32 weeks for women in the IVF group was 2.0 (95% CI [0.9 – 4.1]) when compared to the SMOCC cohort and 0.9 (95% CI [0.7 – 1.1] when compared to the LVR cohort. In terms of number of days, IVF pregnancies were 2.4 days shorter than SMOCC-pregnancies and 2.3 days shorter than LVR-pregnancies, after correcting for maternal and pregnancy variables in the model.

The hypothesis was tested that the effect of IVF exposure may be different for different indications. It is often argued that it is not IVF per se but rather the infertility of the woman and the underlying maternal pathology that causes negative outcomes in IVF pregnancies. If this is the case, women with unexplained infertility are more likely to deliver preterm than women with blocked tubes who are 'fertile' but for the problem with the pathway. Also, women with subfertile partners or with a combination of indications

for treatment would be more likely to deliver prematurely than women with 'pure' tubal pathology. However, when within the IVF group multivariate analyses were carried out calculating the risk of preterm birth before 37 weeks as well as mean pregnancy duration in days, for four separate indication categories while controlling for other factors, differences were found just reaching statistical significance, but in the direction of slightly better outcomes for categories other than tubal pathology. This suggests that the woman's unexplained infertility and the pathology related to it cannot offer the explanation for the poorer results in IVF pregnancies compared to normally conceived pregnancies.

IVF pregnancies have a 5.5 times increased risk of prematurity in our study. IVF singleton pregnancies are approximately 1.5 times more likely to end prematurely and are approximately 2.3 days shorter in duration after taking into account maternal variables, pregnancy variables, iatrogenic preterm births and indication for treatment.

Chapter 4: low birth weight

A group of liveborn, singleton IVF pregnancies and deliveries were studied in terms of their risk of low birth weight for gestational age. Low birth weight < p10 (Small for Gestational Age, SGA) occurred more often in the IVF cohort, 13.8% vs. 9.7% in the SMOCC-group. Low birth weight < p2.3 (Very Small for Gestational age, VSGA) occurred in 3.6% vs. 1.7% in the SMOCC-cohort. Live-born singleton IVF babies had a 1.3% chance of a birth weight < 1500 grams compared to 0.4% in the SMOCC-group. Their odds of birth weight below the 10th percentile was 1.5 (95% CI [1.2-1.9]) compared to the SMOCC control group. Mean birth weight of IVF babies was 186 grams lower than that of control children. After multivariate analysis, controlling for maternal age and height, education, parity, previous miscarriages, pre-existing diabetes and hypertension, smoking and alcohol use as well as infant sex and gestational age, the odds of birth weight below the 10th percentile was 1.4 (95% CI [1.1-1.8]) and the difference in birth weight was 90 grams (p=0.04). Comparing cases with tubal infertility to those in other categories did not change the conclusions. After taking into account a large number of variables including gestational age, a statistically significant difference in birth weight remained. This finding indicates that IVF babies are more likely to be growth restricted compared to control babies even when differences in maternal characteristics and indication for IVF treatment are controlled for. It further indicates that the difference in birth weight cannot be ascribed solely to the increase in preterm birth rate in IVF children.

Chapter 5: outcomes of IVF pregnancies after cryopreservation

For this study the same cohort of IVF births was used as described in the previous chapters. 2460 IVF children were included, 99 of which were born from embryo's that were previously frozen and subsequently thawed and 2361 after cycles with fresh embryo transfer. Multiple pregnancy rate as well as maternal age and height were statistically significantly lower in the "cryo' than in the "fresh" group. The groups did not differ in the other characteristics. Children born after cryopreservation had half the risk of low

birth weight (OR=0.5, 95%CI [0.3-0.9]) and weighed 234 grams more (95% CI [92-377]) than children born after fresh embryo transfer, after correcting for the difference in multiple pregnancy rate and maternal age and height. The difference in risk of growth retardation < P10 between the groups just ceased to be statistically significant after correction (OR=0.5, 95% CI [0.3-1.04])). No difference in gestational age could be observed between the groups after correction. In 89% of cryopregnancies no hyperstimulation was used. In the pregnancies after fresh embryo transfer, no hyperstimulation was used in 4% of preceding cycles. The results indicate that children born after cryopreservation have better perinatal outcomes than those born after IVF with transfer of fresh embryos.

Chapter 6: development of IVF children at two years of age

A total of 197 IVF children and 194 control children were examined with the Bayley test. IVF mothers less often than control mothers had completed higher education (34% vs. 45%), they were less likely to be single parents (1% vs. 5%) and were more likely to take full time care of the IVF child (35% vs. 24%). Thirteen IVF children were born before 36 weeks of gestation, 4 of the controls were. Correction was made in the analyses for differences in parity, age of the mother, education of the mother, single parenthood and caretaking situation. The average test scores for IVF- and control children on the mental and motor developmental index (DI) and on their mental and motor developmental age (DA) with and without correction were the same in the IVF- and the control group. Both the IVF- and control group scored higher than the norm. On the mental scale, no differences in K-scores lower than 5 could be observed between the IVF- and the control group. A K-score lower than 5 indicates clinically significant developmental problems. On the motor score 3.0% (n=6) of IVF children had a score lower than 5 compared to 0% in the control group, a statistically significant difference. Half (n=3) of the low-scoring children were born before 36 weeks of pregnancy and the average duration of pregnancy was 3 weeks lower in this subgroup than in IVF children with normal K-scores. The results of our analysis suggest that IVF children develop normally and that the negative effect on motor development is caused primarily by the excess of preterm births in the IVF group compared to the controls.

Chapter 7: couples' perception of their chance of conception

During 1997 and 1998, men and women in 493 couples who were about to enter into the treatment programme and had not previously undergone IVF were asked, among other questions, what they felt their chance was of having a baby through IVF (duration of treatment or number of cycles anticipated was not specified). The average estimated chance of having a baby through IVF was 44.7%. Among respondents, 30% (n=224) estimated their chance to be exactly 50%, 0.7% (n=5) to be 0%, 4.3% (n=32) to be 100% and 8.1% (n=61) estimated their chance to be higher than 80%. We related respondents' estimates of success to their sex, to duration of their infertility (defined as time between first trying to become pregnant and day of filling out the questionnaire), to type of infertility (tubal only, male subfertility, idiopathic, or a combination) and to age of the female partner. We used a linear regression model, taking into account all factors

simultaneously. Age of the woman and type of infertility appeared to influence the estimate of success, while duration of infertility did not. In the Netherlands, estimated success after three treatments is a useful outcome measure since the first three IVF treatments are paid for by health insurance. This percentage success is claimed by clinicians to be around 50%, which is the most frequently reported respondents' estimate as well. Overall, Dutch couples have expectations of their chances of having a baby after IVF in accordance with physicians' estimates.

Chapter 8: health-related quality of life of couples planning to undergo IVF treatment

Infertility and IVF treatment have been reported to negatively influence the psychosocial status of patients. Suboptimal psychological functioning may lead to discontinuation of the IVF treatment. It is unclear whether certain subgroups may have more problems and if so, whether these problems occur in specific domains. We employed the concept of health-related quality of life to study in which domains people planning IVF may report more problems than the general population does. The concept of health-related quality of life includes four domains: emotional, cognitive, social and physical functioning.

Couples about to enter into IVF treatment, were recruited from six IVF clinics in the Netherlands between January 1996 and June 1997. Participants completed a set of questionnaires. These were the Hopkins Symptom Checklist (HSCL), the Sickness Impact Profile (SIP), the Irrational Beliefs Inventory (IBI) and the Irrational Parenthood Cognitions (IPC-) Scale. A total of 425 men and 447 returned the questionnaire The scores of the IVF group were compared with normative data from the general population. As a group, IVF patients do not have substantially lower health-related quality of life than the general population. However, our results indicate that IVF women, and especially those aged between 21 and 30 years have more social and emotional problems than women of comparable age in the general population. The group of IVF men aged 21 to 30 years report more social and emotional problems as well, compared to men the same age in the general population. We found no differences in social and emotional functioning between the groups of relatively older IVF men and the general population. Further, there were no substantial differences in cognitive and physical functioning for all age groups of IVF men and women compared to the general population.

IVF women of the younger age group had a lower educational level and reported more irrational parenthood cognitions, indicating they have a stronger notion that they need to have children in order to live a happy life. This intense focus on having a child was not found in the older IVF women. Short cognitive counselling interventions may be warranted to enhance younger IVF couples' quality of life.

Conclusions

Twenty two years have passed since the birth of the first IVF baby. First considered to be an unbelievable breakthrough in infertility research, IVF has now become routine treatment. In the Netherlands approximately 1.5% of babies are presently being born after

IVF. Some of the options available thanks to the technology of IVF such as pre-implantation genetic diagnosis, ICSI and post-menopausal motherhood may have seemed sheer medical science fiction in the early days of IVF. Other options seem at the start of the new century at first glance to belong to the distant future but are in fact just around the corner. Storage of ovarian tissue to enable older women to have a child that is biologically and genetically theirs is an example of such a development. A second example is cloning. The first successful cloning of a sheep in 1997 gave rise to speculations about cloning of human beings with the help of IVF. Most governments in industrialised countries have immediately taken position against cloning, but some social commentators have begun to observe that it does not appear to be as objectionable nowadays as it did when it was first introduced to the world (Heitman 1999). One would expect that given the current breadth of practice in IVF many of the medical uncertainties and problems that surrounded it's introduction have by now been solved. Unfortunately, such a conclusion would be premature. A number of problems have become clear yet remain unsolved. Other areas still need a substantial research effort before basic questions will be answered sufficiently.

Our study shows that IVF children are more likely to be born preterm and to be growth retarded. For the largest part, this excess risk is due to the high percentage of multiple pregnancies. This finding has also been shown clearly in many other studies into IVF pregnancies. It can be concluded from our data that as compared to the general population, IVF children are approximately five times more likely to be born prematurely and to be of low birth weight. To prevent higher order multiple pregnancies, Dutch IVF clinics follow a policy of transfer of two embryo's unless maternal age is high and/or the quality of the embryo's is poor, in which case more than two can be transferred. However, since the technique is becoming more successful, the pregnancy rate per cycle as well as the percentage of twin pregnancies appear to be increasing (Te Velde, personal communication). Based on our results and those of other studies it may be concluded that experiments should urgently be started to study the chance of pregnancy as well as the perinatal outcomes with a policy of one embryo per cycle being transferred, in selected subgroups of women that also have good quality embryo's.

Our results show that in singleton IVF pregnancies also, a 1.5 times increased risk occurs of preterm delivery and a two times increased risk of low birth weight and small for gestational age, before taking into account maternal characteristics such as age and parity. These outcomes are in accordance with the existing literature. Clinicians, policy makers and prospective IVF patients need to be aware of this increase and use it in their decision making.
It has been suggested that transfer of fewer embryo's may not only diminish the risk of multiple pregnancies, but also improve the course of singleton IVF pregnancies (Keirse 1995). Indeed some data suggest that preterm birth and low birth weight may be less frequent in singleton pregnancies that result from transfer of a smaller number of embryo's (Doyle 1992). This may be related to the quality of the embryo's. If larger numbers are being transferred it may indicate that the quality of each embryo was perceived by the embryologist as poorer, hence the chance of conception as lower. On the other hand, it may be related to aspects of the initial stages of development in a

pregnancy and indicate it is disadvantageous for a transferred embryo to develop in the company of other embryo's. Whatever may be the case, if experimental policies with one embryo transfer will in the future be studied, it should be researched simultaneously whether IVF singleton pregnancies that result from such a policy, show better outcomes than those that result from multiple embryo transfer in a comparable group.

The increase in risk of adverse outcomes between IVF- and control mothers cannot be explained solely by characteristics of IVF mothers, such as age and parity. In our study, we were able to control in the analysis for a larger number of covariates than most other authors have. Even after correcting for these factors, a statistically significant excess risk of preterm birth (OR=1.6 for SMOCC and OR=1.2 for LVR comparisons), and SGA (OR=1.4) remained. Restricting our analyses to a smaller, low-risk subgroup from the IVF- and control cohorts did not change our conclusion. Our results are in accordance with those of two other recent studies. Bergh et al (1999) studied a large cohort of IVF pregnancies and found an increase in preterm birth and low birth weight in the IVF compared to the control group that is similar to our estimates. In their analysis, stratification was performed by multiplicity of the birth, parity, age, year of delivery and duration of infertility. Koudstaal et al (2000a) find similar outcomes to ours in a cohort of 307 IVF women pair-wise matched to control women on age, height, Body Mass Index, smoking, alcohol consumption, ethnicity and obstetrical/medical history. Similar results can be shown in our analysis of IVF twins compared to control twins, matched on zygosity as well as on a number of other covariates (data to be published). IVF twins show a 1.4 and 1.3 times increased risk of low birthweight and preterm birth, respectively. Koudstaal et al (2000b) compared 96 IVF twins to a similar number of control twins, matched on zygosity and on a large number of demographic and pregnancy variables. IVF twins showed to be more likely to be born prematurely and to be of low birth weight.

It is possible that some aspect of the fertility history or the mother's profile that has not been properly measured, explains this remaining increased risk. The alternative hypothesis is that the IVF procedure itself increases the risk of adverse outcomes. If indeed the IVF procedure increases the risk, it remains unclear as yet, which part of the procedure is most likely to do so. It is suggested from our study that cryo-pregnancies have better perinatal outcomes than regular IVF pregnancies. This finding may point toward the hyperstimulation in the IVF procedure adversely affecting the health of the embryo and the course of the pregnancy. A related explanation may be that the endometrium in unstimulated cycles is of better quality than in stimulated cycles, which may positively influence the growth of the embryo. Most cryo-IVF pregnancies in our study resulted from unstimulated cycles, while no hyperstimulation was used in only 4% of regular IVF cycles. Ovarian hyperstimulation has been reported to increase the circulating relaxin levels (Johnson 1991), while high serum relaxin concentrations have been found to be correllated to preterm birth (Weiss 1993) as well as to number of follicles in the preceding treatment cycle (Kristianson 1996). Olivennes et al (1993) found no difference in preterm birth rate between a group of 162 IVF patients and 263 infertile controls treated with ovarian hyperstimulation. Some studies have suggested that superovulation may lead to low birthweight and growth retardation through elevation of

IGFBP-1 (Insuline Growth Factor Binding Protein) (Lino 1986, Howell 1989, Wang 1991, Johnson 1993, Johnson 1995). Another study suggests that a higher incidence of abnormal placental shapes exists in IVF singleton pregnancies (22% vs 6% in the control group) and that abnormal umbilical cord insertions are more often found (Jauniaux 1990). In assisted reproduction in cattle and sheep, ovulation induction as well as embryo culture are known to lead to larger sized offspring, which in this species is a sign of pathology, as well as to more birth defects and higher perinatal mortality (Willadsen 1991, Walker 1992, Kruip 1997). Outcomes are worse in ovulation induction than in natural conception, yet worse in ovulation induction followed by IVF (Willadsen 1991, Wagtendonk 2000). In animal assisted reproduction, neither the animals donating the gametes, nor the recipients suffer from infertility. Therefore these adverse effects, if real, have to be related to aspects of the procedure itself. Obviously, more research is needed into different aspects of the IVF procedure and their potential adverse effects before specific effects of the IVF procedure or the risk profile of the mothers can be pinpointed.

Collaboration should be sought with animal researchers since the circumstances under which animal IVF take place are much easier to control and to study. This collaboration may also shed some light on the issue of whether or not IVF may cause a higher rate of birth defects in humans. When it becomes clear from animal IVF studies which specific birth defects may be more prevalent, efforts should be mounted to internationally combine human IVF data on birth defects. Only then the numbers may be large enough to allow conclusions about the prevalence of defects in humans equivalent to the ones prominent in cattle and sheep.

In our study, the development of IVF children at two years of age is within normal limits, although an excess risk of abnormal development exists when IVF children are compared to normal controls. This risk appears to be explained by the excess in prematurity. Whether this risk can still be found in our IVF cohort at an older age, remains to be determined. Longer-term risk cannot be excluded since many of the problems related to adverse perinatal circumstances become apparent only at an older age. Very few longer-term studies have been carried out world wide and not enough data are presently available to exclude the chance of long-term problems.

For women treated with IVF also, many potential long term side-effects have as yet been insufficiently researched. A large Dutch study is presently being conducted (van Leeuwen, written communication) to study the possible relationship between IVF treatment and ovarian cancer. Until researched properly it cannot be excluded that other forms of cancer may be more prevalent in women treated with IVF and hyperstimulation.

One very crucial issue that remains unsolved is the issue of the effectiveness of IVF. It was introduced as a treatment option for couples with virtually no chance of spontaneous conception. Obviously, any pregnancy after IVF in a group of women with irreparably damaged, occluded or absent fallopian tubes, may be ascribed to the procedure. At present, however, only one quarter to one third of all IVF treatments are being performed in women with 'pure' tubal pathology. In other words, the remaining couples being treated with IVF have a chance of spontaneous conception that is higher than zero.

Unfortunately, very few studies have compared groups of treated and untreated couples that were alike in all important aspects but in the treatment they received.

To date, only three randomised controlled trials have been carried out comparing IVF treatment to a period of delay with other treatments (Jarell 1993, Soliman 1993, Goverde 1999). The first two studies containing 400 and 200 study subjects respectively, compared IVF to 'other treatments' during a six month waiting period. They were able to show effectiveness of IVF for tubal indications only. The third study compared 86 couples offered intrauterine insemination (IUI) alone to 85 couples offered IUI plus ovarian hyperstimulation to 87 couples offered IVF, with a maximum of 6 cycles in each group. The pregnancy rate per cycle was higher in the IVF than in the control groups. The cumulative pregnancy rate was not higher in the IVF group since couples were more likely to give up treatment before their maximum of six attempts. No trials have yet been carried out comparing IVF to no treatment. Given the fact that couples affected by idiopathic subfertility or male subfertility have an estimated spontaneous conception rate of 2% per cycle (Goverde 1999), it remains to be seen whether IVF can be proven to be effective enough. Observational studies comparing IVF to control groups all suffer from obvious or potential lack of comparability of the treatment and control groups. They cannot provide the necessary answers. In other words, in spite of it's widespread use and the substantial costs accompanying the procedure, IVF treatment is by no means evidence based (Buitendijk 1995). The question whether this technology is more beneficial than conventional treatment or no treatment at all remains unanswered to date. Given that over two decades have passed since the introduction of IVF, many adverse effects have been shown and many potential adverse effects remain, this is a rather astonishing conclusion.

Therefore, a well designed, large, multicenter and possibly multicountry trial should be mounted urgently to show whether or not IVF is effective and if yes, for what indications. In case IVF does appear to be effective for all indications or a subgroup, the perinatal results of this trial can be used to study potential side effects. In a randomised trial, the treated and non-treated group will be far more comparable than will ever be the case in observational studies. The group of children should be followed closely and be studied at older ages for potential developmental problems and ideally, into their own procreative years. Louise Brown's sister, also conceived by IVF, became a teenage mother in 1999. She clearly does not appear to suffer from fertility problems. One case, however, cannot disprove that through the use of fertility techniques medical science is creating a new generation of men and women that may be in need of assisted reproduction themselves. Until a large trial has been mounted, policy makers and physicians should be prudent with the use of IVF.

One may even advocate that couples should not be offered the treatment option unless they consent to be part of a trial. This may certainly be feasible if in one treatment arm the waiting period is regular, while in the other it is shorter. One may argue that it is unethical to introduce uncertainty about the effectiveness of an established procedure that offers hope to many couples with fertility problems. However, infertile couples have in our study shown to be very capable of internalising the information about chances of success as offered to them by their fertility specialist. They appear to be realistically

aware of lower chances with advancing maternal age and type of infertility. Further, younger infertile women appear from our study to have problems interacting socially and appear to have lower quality of life. The message that their chances of spontaneous conception (in case of idiopathic infertility) may be larger than they think, may be beneficial instead of burdening.

Finally, to continue to perform a procedure that is expensive and emotionally challenging to the women that undergo it, while doubts exists with respect to it's effectiveness and while issues of risk of multiple pregnancy, of preterm birth and SGA in singleton pregnancies and developmental problems later in life have not been resolved, inevitably is the most unethical venue. As IVF enters into the new century the time may be ripe for appropriate and more comprehensive studies that take into account issues of effectiveness as well economic, psychological, and short and long term medical aspects of the procedure. Only with the results of those studies will we be able to determine whether IVF has kept the promise it appeared to make 22 years ago. Whether, indeed, it can create the possibility for infertile couples to have a healthy child in a way that is ethically and societally acceptable.

References

Bergh T, Ericson A, Hillensjö T, Nugren KG, Wennerholm UB. Deliveries and children born after in-vitro-fertilisation in Sweden 1982-95: a retrospective cohort study. The Lancet 1999;354:1579-85.

Buitendijk SE. Evidence-based in-vitro fertilisation. The Lancet 1995; 346: 901 (letter).

Edwards R, Steptoe P. A matter of life. 1980. Finestride Ltd. and Crownchime Ltd.

Goverde AJ, McDonnell J, Vermeiden JP, Schats R, Rutten FF, Schoemaker J. Intra-uterine insemination or in-vitro fertilisation in idiopathic subfertility and male subfertility: a randomised trial and cost-effectiveness analysis. Lancet 2000;355(9197):13-8.

Heitman E. Social and ethical aspects of in vitro fertilization. Intl J of Techn Ass in Health Care 1999; 15:1, 22-35.

Howell RJS, Economides D, Teisner B, Farkas AG, Chard T. Placental proteins 12 and 14 in pre-eclampsia. Acta Obstetricia et Gynecologica Scandinavica, 1989;68:237-40.

Jarrell JF, Labell R, Goeree R, Milner R, Collins J. In vitro fertilization and embryo transfer: A randomized controlled trial. The Online Journal of Current Clinical Trials, Document no.73, July 1993.

Jauniaux E, Englert Y, Vanesse M, Hiden M, Wilkin P. Pathologic Features of Placentas From Singleton Pregnancies Obtained by In Vitro Fertilization and Embryo Transfer. Obstetrics and Gynecology, 1990;76:61-4.

Johnson MR, Okokon E, Collins WP, Sharma V, Lightman SL. The effect of human chorionic gonadotropin and pregnancy on the circulating level of relaxin. J Clin Endocr Metabol 1991;72:1042-7.

Johnson MR, Abbas A, Norman-Taylor JQ, et al. Circulating placental protein 14 (PP14) in the first trimester of spontaneous and IVF pregnancies. Human Reproduction, 1993;8:323-6.

Johnson MR, Irvine R, Hills F, et al. Superovulation, IGFBP-1 and birthweight. European Journal of Obstetrics & Gynaecology and Reproductive Biology, 1995;59:193-5.

Keirse MJNC, Helmerhorst FM. The impact of assisted reproduction on perinatal health care. Soz Präventivmed 1995;40:343-51.

Koudstaal N, Braat DDM, Bruinse HW, Naaktgeboren N, Vermeiden JPW, Visser GHA. Obstetric outcome of singleton pregnancies after in vitro fertilization. A pair-wise matched cohort study in four Dutch University hospitals. Accepted for publication, Hum Reprod. (2000a).

Koudstaal J, Bruinse HW, Helmerhorst FM, Vermeiden JPW, Willemsen WNP, Visser GHA. Obstetric outcome of twin pregnancies after in-vitro fertilization: a matched control study in four Dutch University hospitals. Hum Reprod 2000b;15(4):935-40.

Kristiansson P, Svärdsudd, Schoultz B von, Wramsby H. Supraphysiological serum relaxin concentration during pregnancy achieved by in-vitro fertilization is strongly correlated to the number of growing follicles in the treatment cycle. Hum Reprod 1996:11:2036-40.

Kruip ThAM, Daas JHG den. In vitro produced and cloned embryos: effects on pregnancy, parturition and offspring. Theriogenology 1997;47:43-52.

Lino K, Sjoberg J, Seppala M. Elevated circulating levels of a decidual protein, placental protein 12, in preeclampsia. Obstetrics and Gynecology, 1986;68:58-60.

Olivennes F, Rufat P, André B, Pourade A, Quiros MC, Frydman R. The increased risk of complication observed in singleton pregnancies resulting from in-vitro fertilization (IVF) does not seem to be related to the IVF method itself. Hum Reprod 1993;8:1297-1300.

Soliman S, Daya S, Collins J, Jarrell J. A randomized trial of in vitro fertilization versus conventional treatment for infertility. Fertility and Sterility, 1993; 59:1239-44.

Wagtendonk van-de Leeuw AM, Mullaart E, Roos de APW, Merton JS, Daas den JHG, Kemp B, Ruigh de L. Effects of different reproduction techniques: AI, MOET or IVP, on health and welfare of bovine offspring. Theriogenology 2000; 53:575-597.

Walker SK, Heard TM, Seamark RF. In vitro culture of sheep embryos without co-culture: Sucesses and perspectives. Theriogenology, 1992; 37:111-26.

Wang HS, Lim J, English J, Irvine L, Chard T. The concentration of insulin-like growth factor (IGF-1) and insulin-like growth factor-binding protein-1 (IGFBP-1) in human umbilical cord serum at delivery: relation to fetal weight. Journal of Endocrinology, 1991;129:459-64.

Weiss G, Goldsmith LT, Sachdev R, Hagen S von, Lederer K. Elevated first-trimester serum relaxin concentrations in pregnant women following ovarian stimulation predict prematurity risk and preterm delivery. Obstet Gynecol 1993;82:821-8.

Willadsen WM, Janzen RE, McAlister RJ et al. The viability of late morulae and blastocysts produced by nuclear transplantation in cattle. Theriogenology, 1991; 35:161-170.

Chapter 9

Samenvatting en bespiegelende conclusies

Inleiding

In vitro fertilisatie (IVF) staat op een prominente plaats in de vele overzichten van belangrijke medische ontwikkelingen in de vorige eeuw. Samen met de komst van antibiotica, medische diagnostiek zoals ultrageluid en MRI, immunisatieprogramma's, bloedtransfusies en het in kaart brengen van het menselijke genoom staat IVF bovenaan de lijst met mijlpalen onder de ontdekkingen die het aanzien van de geneeskunde hebben veranderd. En terecht. Tot de introductie van IVF hadden specialisten weinig mogelijkheden om de vruchtbaarheid te bevorderen. Opwekking van de ovulatie, chirurgische ingrepen aan de eileider en kunstmatige inseminatie boden mogelijkheden voor echtparen die moeilijk in verwachting raakten, maar dat gold slechts voor een klein deel van hen. Sinds IVF en ICSI op grote schaal beschikbaar kwamen kan vrijwel elk paar met een vruchtbaarheidsprobleem de hoop worden geboden dat ze een biologisch kind van zichzelf kunnen krijgen. Met betrekking tot de sociale, ethische, wettelijke en medische consequenties, kan IVF terecht de meest verstrekkende medische ontwikkeling van de 20e eeuw worden genoemd, waarbij het wellicht die positie alleen hoeft te delen met het menselijke-genoomproject.

Toch is IVF destijds ontwikkeld – zoals geldt voor de meeste huidige medische technologieën – zonder dat veel aandacht werd besteed aan de sociale, ethische, wettelijke en medische gevolgen. De bedenkers van IVF, Edwards en Steptoe, beschrijven in hun boek 'A matter of life' verbazingwekkend gedetailleerd wat zij doormaakten, zowel privé als beroepsmatig, bij de ontwikkeling van de techniek en uiteindelijk de realisatie van de geboorte van Louise Brown (Ewards & Steptoe, 1980). Daarentegen besteden zij weinig aandacht aan de mogelijke sociale gevolgen van hun uitvinding. Edwards beschrijft de nachten die hij in zijn laboratorium doorbrengt tijdens zijn pogingen om met zijn eigen sperma eicellen te bevruchten die hij van bevriende gynaecologen had gekregen. Hij ontving die eicellen in ovariumweefsel dat overbleef na chirurgische ingrepen. Hij beschrijft zijn strijd bij de ontwikkeling van een kweekmedium waarin sperma kan worden klaargemaakt om eicellen te bevruchten. In zijn pogingen om dat te bereiken ontwikkelde hij een semipermeabel doosje waarin spermacellen konden worden opgeslagen. Dit doosje werd vervolgens ingebracht in de baarmoeders van patiënten die in een ziekenhuis wachtten op een gynaecologische chirurgische behandeling. Hij beschrijft zijn vele pogingen om een kweekmedium te ontwikkelen om de embryo's tot ontwikkeling te brengen. Zijn pogingen maakten een geweldige sprong voorwaarts toen hij op de hoogte kwam van de vorderingen van Steptoe met laparoscopische chirurgische technieken. En samen beschrijven Edwards en Steptoe hoe, vele jaren later, hun inspanningen werden bekroond met de zwangerschap van mevrouw Brown en uiteindelijk de geboorte van Louise in 1978. Terwijl de auteurs werken aan de ontwikkeling en perfectionering van de techniek lijken zij zich weinig zorgen te maken over mogelijke bijwerkingen van de procedure. Edwards en Steptoe verklaren dat vier kinderartsen na de geboorte van Louise haar in totaal aan 41 tests hebben onderworpen, waarna zij tot de conclusie kwamen dat zijn een volkomen normale baby was, 'wat wij al wisten'. Wettelijke, ethische en sociale bedenkingen over de introductie van de procedure worden wel in hun boek genoemd, maar voornamelijk als vervelende aspecten die hun pogingen de techniek breder te laten toepassen in de weg

stonden. Toch had de ontwikkeling van IVF verreikende gevolgen voor de voortplanting. IVF veranderde de maatschappelijke en medische betekenis van ouderschap en leverde ongekende complicaties op voor beleidsmakers.

In de twintig jaar na de introductie van IVF is veel discussie gevoerd over de medicalisering van zwangerschap en voortplanting ten gevolge van IVF. Edwards en Steptoe hadden IVF ontwikkeld om te worden toegepast bij een beperkt aantal indicaties, vooral verband houdend met slechte doorgankelijkheid of afwezigheid van eileiders. Het aanvankelijke succes van IVF leidde al gauw tot verbreding van het aantal toepassingen. De verruiming van de toepassing van IVF leidde tot een snel stijgende vraag naar deze behandeling en daardoor tot tal van wettelijke en maatschappelijke problemen: het probleem van overtollige embryo's alsmede hun opslag en eigendomsrechten; het probleem van donoreicellen en de definitie van biologische grenzen aan de vruchtbare leeftijd; het probleem van draagmoederschap; de vraag of embryo's al of niet voor medisch onderzoek mogen worden gebruikt. De wetgeving in de meeste landen bleek ontoereikend om veel van deze vraagstukken op te lossen, wat blijkt uit de grote verschillen in nationale regelgeving. Bovendien deden zich zo snel nieuwe ontwikkelingen rond IVF voor dat het vrijwel onvermijdelijk was dat wet- en regelgeving sterk achter de feiten aan bleef lopen.

De nadruk op complexe wettelijke en ethische vraagstukken gecombineerd met het accent van de medische stand op verbetering van de effectiviteit van de procedure heeft er in de eerste jaren na de verspreiding van de IVF-techniek toe geleid dat betrekkelijk weinig aandacht werd besteed aan het verloop en de uitkomst van IVF-zwangerschappen. Elke afzonderlijke fase van de IVF-behandeling zou het ontwikkelende embryo kunnen schaden. Ten eerste zijn de meeste hormoonpreparaten die bij de IVF-behandeling worden gebruikt niet uitvoerig getest op mogelijk teratogene effecten. Ten tweede is het nog onduidelijk of de verwijdering van de eicel tot mechanische schade aan die cel leidt. Ten derde worden bevruchte eicellen in een kweekmedium overgebracht om ze verder te laten groeien; het kan niet worden uitgesloten dat bestanddelen uit het medium negatieve effecten op het embryo kunnen hebben. Ten vierde zou de procedure voor het overbrengen van het embryo in de baarmoeder na hun ontwikkeling schade kunnen aanrichten. En ten slotte zou het invriezen van overtollige embryo's tot mechanische beschadiging kunnen leiden. De mogelijke negatieve effecten in één of meer van deze fases kunnen leiden tot een suboptimale ontwikkeling van het embryo. Bovendien komt het embryo op een onnatuurlijke wijze in de baarmoeder, en mogelijk ook niet in het optimale stadium van zijn ontwikkeling, terwijl het baarmoederslijmvlies zich onder abnormale, kunstmatig-gestimuleerde omstandigheden heeft ontwikkeld. Elk van deze aspecten kan op zichzelf leiden tot een abnormale ontwikkeling van het embryo, die vervolgens kan leiden tot perinatale problemen zoals groeiachterstand, vroeggeboorte of aangeboren afwijkingen.

In de eerste jaren na de introductie van IVF was de psychologische toestand van paren die een IVF-behandeling ondergingen evenmin een onderwerp waarnaar veel onderzoek werd gedaan. Er werd kennelijk van uitgegaan dat onvrijwillige kinderloosheid

vanzelfsprekend een veel grotere psychologische last betekent dan een behandeling om die kinderloosheid op te lossen ooit met zich mee zou kunnen brengen.

Hoofdstuk 2: Overzicht van de literatuur over de uitkomst van zwangerschappen en de ontwikkeling van het kind

In hoofdstuk 2 wordt een overzicht gegeven van onderzoeksgegevens gepubliceerd tot begin 1999 over de gezondheid van kinderen geboren na een IVF-behandeling. Aan het eind van de jaren '90 zijn een aantal artikelen over dat onderwerp verschenen. Een van de grootste problemen bij IVF is het hoge percentage meerlingen. Van alle IVF-zwangerschappen betreffen 25–30% meerlingen, inclusief 3% gevallen waarin de vrouw zwanger is van drie of meer kinderen. Bij bijna de helft van de kinderen geboren na een IVF-behandeling is dus sprake van meerlingen. Meerlingzwangerschap brengt het inherente risico op complicaties met zich mee, vooral vroeggeboorte, vertraagde groei in de baarmoeder en perinatale sterfte. Maar ook eenlingen geboren na een IVF-behandeling blijken een verhoogd risico op negatieve uitkomsten te hebben. Analyse van de literatuur leert dat bij IVF-eenlingen een ongeveer twee keer zo hoge kans op vroegtijdige geboorte en een laag geboortegewicht als kinderen die op een natuurlijke manier zijn verwekt. Uit dit onderzoek kan niet worden geconcludeerd of dit verhoogde risico samenhangt met de IVF-procedure zelf of met factoren die te maken hebben met het profiel van de moeders. Moeders die bevallen na een IVF-behandeling zijn gemiddeld ouder en zijn vaker voor de eerste keer zwanger dan andere moeders. Zij hebben mogelijk ook een hogere gemiddelde sociaal-economische status en kunnen verschillen in riskante levensstijl (zoals roken en drinken) tijdens de zwangerschap. Zij lijden mogelijk ook vaker aan bepaalde chronische aandoeningen die niet alleen de vruchtbaarheid maar ook de uitkomst van een zwangerschap beïnvloeden. De onvruchtbaarheid op zich, in plaats van de IVF-behandeling, zou dus een negatief effect op de uikomst kunnen hebben. Geen van de onderzoeken gevonden in de literatuur tot medio 1999 hebben de uitkomsten tegelijkertijd gecorrigeerd voor een groot aantal verstorende variabelen, waardoor het onmogelijk is conclusies te trekken over de oorzaak van het verhoogde risico. Over het geheel bezien is het percentage meerlingen echter de belangrijkste risicofactor voor een ongunstige uitkomst bij een IVF-zwangerschap.

Een aantal studies is gewijd aan afwijkingen bij de geboorte na een IVF-behandeling. Geen van deze studies vermeldt een verhoogd percentage afwijkingen bij IVF-kinderen. Om een mogelijk verhoogd risico op specifieke, goed gedefinieerde afwijkingen te kunnen vaststellen, zijn echter grotere aantallen kinderen en pooling van de analyses nodig.

In een klein aantal studies is gekeken naar de mogelijke effecten van het invriezen van embryo's. Invriezen en ontdooien leiden tot belangrijke cellulaire veranderingen. Het is niet duidelijke of deze veranderingen een negatieve invloed hebben op het nageslacht. Tot dusver uitgevoerd onderzoek heeft geen negatieve uitkomsten aangetoond, maar de aantallen kinderen betrokken in dit onderzoek is te klein om daar erg zeker van te zijn.

Slechts weinig onderzoek is gewijd aan de ontwikkeling van het kind na een IVF-zwangerschap. Bovendien was de groep kinderen betrokken in dat onderzoek relatief klein en werden de kinderen over een korte periode gevolgd. Veel aandoeningen, zoals problematische schoolprestaties of ADHD, kunnen pas jaren later worden gediagnostiseerd. Deze literatuur heeft geen negatieve effecten op de ontwikkeling op zichzelf of op de daarmee gepaard gaande vroeggeboorte van IVF-kinderen aangetoond. Sommige studies laten zien dat de kwaliteit van ouderschap wellicht beter is na een IVF-zwangerschap dan in de controlegroep. Als er al negatieve effecten van IVF op de ontwikkeling van het kind bestaan, worden deze gecompenseerd door het feit dat de zwangerschap echt gewenst was en door de hogere kwaliteit van ouderschap. Kortom, uit de literatuur gepubliceerd tot begin 1999 kan niet worden beoordeeld of IVF-kinderen zich normaal ontwikkelen of dat de groepen kinderen betrokken in zulk onderzoek te klein waren om verschillen in ontwikkeling te kunnen vaststellen.

Doelstellingen van het onderzoek

In 1992 werd door Nederlandse IVF-artsen en onderzoekers een onderzoek opgezet met een grote onderzoeksgroep en een geschikte controlegroep die conclusies over de gezondheid van IVF-baby's zou kunnen opleveren. Kinderen geboren uit IVF-zwangerschappen zouden direct na hun geboorte worden bestudeerd en worden vergeleken met en controlegroep. Een subgroep van deze IVF-kinderen zou worden bekeken wanneer zij 2 jaar oud waren en worden vergeleken met een controlegroep wat betreft hun psychomotorische ontwikkeling. In een afzonderlijke groep zouden de effecten van de IVF-procedure op de psychologische status en de kwaliteit van leven van IVF-kinderen worden bestudeerd. De onderzoeksvragen waren:

- verschillen IVF-kinderen van natuurlijk verwerkte kinderen in geboortegewicht en zwangerschapsduur?
- geeft de geboorte van IVF-kinderen na invriezen van het embryo een andere uitkomst dan wanneer het embryo niet is ingevroren?
- verschillen IVF-kinderen van natuurlijk verwekte kinderen in psychomotorische ontwikkeling als ze 2 jaar oud zijn?
- wat is de invloed van IVF op de kwaliteit van leven van paren die deze behandeling ondergaan?

Resultaten van het onderzoek

Hoofdstuk 3: Vroeggeboorte

Het percentage vroeggeboorten werd bestudeerd bij een groep van 2043 IVF-zwangerschappen waaruit 2636 kinderen werden geboren. De resultaten werden vergeleken met twee afzonderlijke controlegroepen, de SMOCK (Sociaal Medisch Onderzoek ConsultatiebureauKinderen)- en de LVR (Landelijke Verloskunde Registratie)-groep. IVF-moeders hadden gemiddeld minder kinderen, hadden een hogere opleiding genoten en maakten minder vaak deel uit van een etnische minderheid dan moeders in de twee controlegroepen. Zij rookten minder en dronken minder alcohol dan moeders in het SMOCK-cohort. Zij leden minder aan diabetes of hoge bloeddruk vóór de

zwangerschap dan vrouwen in het SMOCK-cohort. In de IVF-groep waren de aantallen opgewekte weeën en primaire keizersneden hoger dan in de LVR-groep. Minder IVF-moeders dan vrouwen in de controlegroepen bevielen thuis. Van de IVF-zwangerschappen leidden 26,5% tot meerlingen, tegenover 1,6% in beide controlegroepen. Vroeggeboorten kwamen vaker voor in de IVF-groep (23,8%) dan in de SMOCK-groep (6,0%) of de LVR-groep (7,6%). Meer kinderen in de IVF-groep (5,2%) dan in het LVR-cohort (1,6%) overleden in de perinatale periode. Eenlingen geboren uit een IVF-zwangerschap hadden een grotere kans vroegtijdig te worden geboren dan eenlingen in de controlegroepen (9,3% tegenover 5,2% in de SMOCK-groep en 6,9% in de LVR-groep). Ook overleden meer IVF-eenlingen (3,0%) dan kinderen in het LVR-cohort (1,5%) in de perinatale periode. Logistische regressieanalyse werd uitgevoerd om tegelijkertijd te corrigeren voor een groot aantal variabelen. Na correctie voor leeftijd, pariteit, etnische achtergrond, diabetes en hoge bloeddruk vóór de zwangerschap, opwekking van de weeën en toepassing van keizersnede, was de odds ratio (OR) voor IVF-vrouwen voor beëindiging van de zwangerschap vóór de 37e week 1,6 (95% betrouwbaarheidsinterval (BI) [1,2–2,1]) vergeleken met de SMOCK-groep en 1,2 (95% BI [1,1–1,4]) in vergelijking met het LVR-cohort. De OR voor IVF-vrouwen voor vroeggeboorte vóór de 32e week was 2,0 (95%BI [0,9–4,1]) vergeleken met de SMOCK-groep en 0,9 (95% BI [0,7–1,1]) in vergelijking met het LVR-cohort. IVF-zwangerschappen duurden gemiddeld 2,4 dagen korter dan zwangerschappen in de controlegroepen in het model waarin werd gecorrigeerd voor lengte, pariteit, etnische achtergrond, eerdere miskramen, opleiding en diabetes en hoge bloeddruk vóór de zwangerschap.

De hypothese werd getest dat het effect van de IVF-behandeling verschilt per indicatie. Er wordt vaak gesteld dat het niet IVF op zichzelf is maar de onvruchtbaarheid van de vrouw en de onderliggende pathologie die verantwoordelijk zijn voor een negatieve uitkomst bij een IVF-behandeling. Als dat zo is, kan worden verwacht dat vrouwen met een onverklaarde onvruchtbaarheid een grotere kans hebben voortijdig te bevallen dan vrouwen met geblokkeerde eileiders, die 'vruchtbaar' zijn afgezien van het probleem van de doorgang van de eicel. Ook zouden dan vrouwen met een minder vruchtbare partner of met een combinatie van indicaties voor IVF-behandeling een grotere kans hebben voortijdig te bevallen dan vrouwen met een 'zuiver' eileiderprobleem. De verschillen die werden gevonden in een multivariate analyse binnen de IVF-groep waarin de kans op vroeggeboorte vóór de 37e week en de gemiddelde zwangerschapsduur werd berekend voor vier verschillende indicatiecategorieën, wijzen echter meer de kant uit van iets een verhoogd risico voor vrouwen met tubapathologie. Dit geeft aan dat onverklaarde onvruchtbaarheid van de vrouw en de daaraan gerelateerde pathologie geen verklaring biedt voor de slechtere resultaten van IVF-behandelingen vergeleken met 'normale' zwangerschappen.

Uit ons onderzoek bleek een 5,5 maal zo hoge kans op vroeggeboorte in de IVF-groep dan in de andere groepen. IVF-eenlingzwangerschappen hadden een ca. 1,5 maal hogere kans op vroeggeboorte dan 'normale' zwangerschappen en duurden gemiddeld 2,3 dagen korter na correctie voor variabelen van de moeder, variabelen van de zwangerschap, iatrogene vroeggeboorte en indicaties voor IVF-behandeling.

Hoofdstuk 4: Laag geboortegewicht

Een groep levendgeboren eenlingen uit IVF-zwangerschappen werd bestudeerd met betrekking tot risico op een laag geboortegewicht gecorrigeerd voor zwangerschapsduur. Een laag geboortegewicht (beneden de 10de percentiel, 'small for gestational age', SGA) kwam vaker voor in het IVF-cohort (13,8%) dan in de SMOCK-groep (9,7%). Een zeer laag geboortegewicht (beneden de 2,3^{de} percentiel, 'very small for gestational age', VSGA) kwam voor bij 3,6% in de IVF-groep tegenover 1,7% in de SMOCK-groep. Levendgeboren IVF-eenlingen hadden een kans van 1,3% op een geboortegewicht beneden 1500 gram tegenover 0,4% van de eenlingen in de SMOCK-groep. De OR in die groep voor een geboortegewicht beneden de 10de percentiel was 1,5 (95%BI [1,2–1,9]) vergeleken met de SMOCK-groep. Het gemiddelde geboortegewicht van IVF-baby's was 186 gram lager dan dat van baby's in de controlegroepen. In een multivariate analyse waarin werd gecorrigeerd voor de leeftijd en de lengte van de moeder, opleiding, pariteit, eerder miskramen, diabetes vóór de zwangerschap, roken en alcoholgebruik evenals voor het geslacht van het kind en de zwangerschapsduur, was de OR voor een geboortegewicht beneden de 10de percentiel 1,3 (95%BI [1,1–1,8]) en waren de IVF-kinderen gemiddeld 90 gram lichter (p=0,04). Vergelijking van vrouwen met een eileiderprobleem met andere categorieën leverde geen andere conclusie op.. Na correctie voor een groot aantal variabelen met inbegrip van zwangerschapsduur, bleef er een significant verschil in geboortegewicht. Deze bevinding geeft aan dat IVF-baby's meer kans hebben op een vertraagde groei dan natuurlijk verwekte kinderen, zelfs als rekening wordt gehouden met verschillen in kenmerken van de moeder en indicaties voor IVF-behandeling. Dit wijst er tevens op dat het verschil in geboortegewicht niet louter kan worden toegeschreven aan de hogere kans op vroeggeboorte bij IVF-kinderen.

Hoofdstuk 5: Uitkomsten van IVF-zwangerschappen na het invriezen van embryo's

Voor dit onderzoek werd hetzelfde IVF-cohort gebruikt als in de voorafgaande hoofdstukken beschreven. Het onderzoek omvatte 2460 IVF-kinderen waarvan 99 waren geboren uit embryos's die voorheen waren ingevroren (cryopreservering) en vervolgens ontdooid; de overige 2361 waren ontstaan na een cyclus waarin een of meerdere verse embryo's waren overgebracht. Zowel het percentage meerlingen als de leeftijd en de lengte van de moeder lagen significant hoger in de 'cryo'-groep dan in de 'verse' groep. De groepen verschilden niet in andere kenmerken. Kinderen geboren na cryopreservering hadden minder kans op een laag geboortegewicht (OR 0,5, 95%BI [0,3–0,9]) en wogen gemiddeld 234 gram meer (95%BI [92–377]) dan kinderen geboren na het overbrengen van een verse embryo als de uitkomsten werden gecorrigeerd voor het verschil in kans op meerlingen en de leeftijd en lengte van de moeder. Het verschil in kans op vertraagde groei (beneden de 10de percentiel) tussen de groepen was net niet meer significant na correctie voor deze factoren (OR 0,5, 95%BI [0,3–1,04]). Na correctie voor deze variabelen werd geen verschil in zwangerschapsduur tussen de groepen gevonden. In 89% van de zwangerschappen na cryopreservering was geen hyperstimulatie gebruikt. In de zwangerschappen na het overbrengen van een vers embryo was in 4% van de voorgaande cycli geen hyperstimulatie toegepast. Deze resultaten geven aan dat kinderen

geboren na cryopreservering een betere perinatale uitkomst hebben dan IVF-kinderen geboren na het overbrengen van een vers embryo.

Hoofdstuk 6: De ontwikkeling van IVF-kinderen op een leeftijd van 2 jaar

Een groep van 197 IVF-kinderen en 194 controlekinderen werd onderworpen aan de Bayley-test. Minder IVF-moeders dan moeders in de controlegroep hadden een HBO of universitaire opleiding afgerond (resp. 34% en 45%). Zij waren minder vaak de alleenstaand ouder (resp. 1% en 5%) en namen vaker de volledige zorg voor de kinderen op zich (resp. 35% en 24%). Dertien IVF-kinderen werden vóór de 36e week van de zwangerschap geboren, tegenover vier in de controlegroep. Beide groepen verschilden niet in gemiddelde testwaarden voor de mentale en motorische ontwikkelingsindex ('developmental index', DI) en op hun mentale en motorische ontwikkelingsleeftijd, ongeacht of de analyseresultaten werden gecorrigeerd voor verschillen in pariteit, leeftijd van de moeder, opleiding van de moeder, eenouderschap en zorgsituatie (al of niet de volledige zorg voor het kind). Zowel de IVF-groep als de controlegroep hadden gemiddelde scores boven de norm. Op de mentale ontwikkelingsschaal werden geen verschillen tussen de groepen gevonden in K-scores lager dan 5 (een waarde die duidt op klinisch relevante gedragsproblemen). Op de motorische schaal hadden 6 kinderen (3,0%) in de IVF-groep een score lager dan 5 tegenover geen enkel kind in de controlegroep. Dit verschil was significant. Van die zes kinderen waren drie geboren vóór de 36e week van de zwangerschap. De gemiddelde zwangerschapsduur was bij de groep met de lage K-score 3 weken korter dan bij IVF-kinderen met een normale K-score. De resultaten van de analyse wijzen erop dat IVF-kinderen zich normaal ontwikkelen en dat het negatieve effect van IVF op hun motorische ontwikkeling waarschijnlijk primair valt toe te schrijven aan het hogere percentage vroeggeboorten vergeleken met de controlegroep.

Hoofdstuk 7: Perceptie van vrouwen en mannen van de kans op succes bij IVF-behandeling

In 1997 en 1998 werd 493 paren die voor de eerste keer op het punt stonden met een IVF-behandeling te beginnen onder meer gevraagd hoe groot zij de kans achtten dat zij door middel van IVF een kind zouden krijgen (de behandelingsduur en het aantal cycli dat de behandeling zou duren werd buiten beschouwing gelaten). Gemiddeld over alle respondenten werd die kans geschat op 44,7%: 30% (n = 224) schatte die kans op precies 50%, 0,7% (n = 5) op 0%, 4,3% (n = 32) op 100% en 8,1% (n = 61) op meer dan 80%. De door respondenten geschatte kans op succes werd gerelateerd aan hun geslacht, de duur van hun onvruchtbaarheid (gedefinieerd als de tijd tussen de dag waarop de vrouw voor het eerst probeerde zwanger te raken tot de dag waarop de vragenlijst werd ingevuld), de aard van de onvruchtbaarheid (alleen een eileiderprobleem, verminderde mannelijke vruchtbaarheid, idiopathische onvruchtbaarheid, of een combinatie daarvan) en de leeftijd van de vrouw. Er werd een lineair regressiemodel gebruikt waarin met al deze factoren tegelijk rekening wordt gehouden. De leeftijd van de vrouw en de aard van de onvruchtbaarheid bleken van invloed op de geschatte kans op succes, maar de duur van de onvruchtbaarheid niet. In Nederland is het geschatte succes na drie behandelingen

een zinvolle maat voor resultaat omdat meestal alleen de eerste drie behandelingen worden vergoed door verzekeringsmaatschappijen. Dat succespercentage wordt door artsen geclaimd rond de 50% te liggen, wat ook de gemiddelde schatting door de respondenten was. Over het geheel genomen hebben Nederlandse paren kennelijk verwachtingen ten aanzien van hun kans op een kind door middel van IVF die overeenkomt met de schattingen van artsen.

Hoofdstuk 8: Gezondheidsgerelateerde kwaliteit van leven van paren die van plan zijn een IVF-behandeling te ondergaan

Onvruchtbaarheid en het ondergaan van een IVF-behandeling kunnen het psychosociale functioneren van mensen negatief beinvloeden. Suboptimaal psychosociaal functioneren kan leiden tot het afbreken van de IVF-behandeling. Het was totnogtoe niet duidelijk of bepaalde subgroepen een grotere kans hadden op problemen en zo ja, of die problemen zich dan met name in bepaalde domeinen voordoen. We gebruikten het concept 'gezondheidsgerelateerde kwaliteit van leven' om te onderzoeken in welke domeinen mensen die op het punt staan een IVF-behandeling te ondergaan, mogelijk meer problemen rapporteren dan de algemene bevolking.

Tussen januari 1996 en juni 1997 werden paren met vruchtbaarheidsproblemen die op het punt stonden met en IVF-behandeling te beginnen, geworven bij zes IVF-klinieken. De deelnemers vulden een serie vragenlijsten in, namelijk de Hopkins Symptom Checklist (HSCL), de Sickness Impact Profile (SIP), de Irrational Beliefs Inventory (IBI) en de Irrational Parenthood Cognitions (IPC)-schaal. De vragenlijsten werden teruggestuurd door 425 mannen en 447 vrouwen. De scores van deze IVF-groep werden vergeleken met normgegevens over de totale bevolking. Gemeten over het gehele cohort ondervonden IVF-patiënten geen substantieel lagere gezondheidsgerelateerde kwaliteit van leven. De resultaten gaven echter wel aan dat IVF-vrouwen, vooral die met een leeftijd tussen 21 en 30 jaar, meer sociale en emotionele problemen hadden dan vrouwen van vergelijkbare leeftijd in de totale bevolking. Ook IVF-mannen in de leeftijd van 21–30 jaar meldden meer sociale en emotionele problemen dan mannen van die leeftijd in de totale bevolking. Er werden geen verschillen in sociaal en psychisch functioneren gevonden tussen oudere IVF-mannen en hun leeftijdsgenoten in de totale bevolking. Daarnaast werden voor geen van de leeftijdsgroepen van mannen en vrouwen in het IVF-cohort belangrijke verschillen in cognitief en fysiek functioneren gevonden ten opzichte van de totale bevolking.

Jongere IVF-vrouwen hadden gemiddeld een lagere opleiding dan oudere IVF-vrouwen. Zowel jongere IVF-vrouwen als jongere IVF-mannen gaven vaker irrationele cognities ten aanzien van ouderschap aan, wat erop duidt dat zij een sterker gevoel hebben dat zij kinderen nodig hebben om een gelukkig bestaan te kunnen lijden. Deze intense nadruk op het krijgen van een kind werd niet gevonden bij oudere IVF-vrouwen. Deze IVF-paren zouden gerichte kortdurende therapie aangeboden moeten krijgen om hun subjectieve kwaliteit van leven te verhogen.

Conclusies

Er zijn 22 jaar verstreken sinds de eerste IVF-baby werd geboren. Waar IVF aanvankelijk werd beschouwd als een ongelooflijke doorbraak in het vruchtbaarheidsonderzoek is het nu een routinebehandeling geworden. In Nederland worden thans ca. 1,5% van de kinderen na IVF geboren. Sommige opties die de IVF-technologie nu biedt, zoals genetische diagnostiek vóór het inbrengen van het embryo, ICSI en moederschap na de menopauze, leken louter medische science fiction in de eerste jaren na de introductie van IVF. Nieuwe mogelijkheden lijken aan het begin van deze eeuw op het eerste gezicht iets van de verre toekomst, maar liggen in feite om de hoek. Het invriezen van ovariumweefsel om oudere vrouwen in staat te stellen een kind te krijgen dat biologisch en genetisch van henzelf is, behoort tot dit soort ontwikkelingen. Klonen is een ander voorbeeld. Toen in 1997 voor het eerst met succes een schaap werd gekloond, gaf dat aanleiding tot speculaties over het klonen van mensen met behulp van IVF. In de meeste geïndustrialiseerde landen heeft de regering meteen stelling genomen tegen klonen, maar sommige maatschappelijke commentatoren beginnen waar te nemen dat klonen vandaag de dag niet meer zo verwerpelijk lijkt als vlak na de introductie van die technologie (Heitman 1999). Men zou verwachten dat, gezien de huidige ruime toepassing van IVF, veel van de medische onzekerheden en problemen die de introductie van IVF vergezelden nu wel tot het verleden zullen horen. Deze conclusie is helaas voorbarig. Een aantal problemen is inmiddels duidelijk geworden maar wachten nog op een oplossing. Op andere terreinen is nog veel onderzoek nodig voordat fundamentele vragen naar tevredenheid kunnen worden beantwoord.

Ons onderzoek laat zien dat IVF-kinderen een grotere kans hebben dan natuurlijk verwekte kinderen om voortijdig te worden geboren en vertraging in hun groei op te lopen. Dit verhoogde risico heeft grotendeels te maken met meerlingzwangerschappen. Ook uit andere onderzoeken is dit duidelijk gebleken. Uit onze gegevens kan worden afgeleid dat, vergeleken met de algemene populatie, IVF-kinderen een ongeveer vijf keer zo groot risico hebben op vroeggeboorte en laag geboortegewicht. Om grotere meerlingzwangerschappen (drie of meer) te voorkomen volgen Nederlandse IVF-klinieken een beleid waarbij niet meer dan twee embryo's worden teruggeplaatst, tenzij de leeftijd van de moeder hoog is en/of de kwaliteit van de embryo's heel slecht. Ondanks dit beleid lijkt het percentage meerlingzwangerschappen na IVF-behandeling te stijgen (Te Velde, persoonlijke communicatie), waarschijnlijk omdat de succeskans van de procedure toeneemt en daarmee de kans op een tweelingzwangerschap. Gebaseerd op onze resultaten kan worden geconcludeerd dat dringende onderzocht moet worden wat de kans is op zwangerschap en wat de perinatale uitkomsten zijn bij een beleid van terugplaatsing van slechts een embryo, in een geselecteerde subgroep waarbij tevens embryo's van goede kwaliteit kunnen worden verkregen.

Ons onderzoek laat echter ook zien dat bij IVF-eenlingzwangerschappen een 1,5 maal groter risico bestaat op vroeggeboorte en een twee keer groter risico op laag geboortegewicht en groeivertraging, voordat rekening wordt gehouden met maternale karakteristieken zoals leeftijd en pariteit. Deze bevindingen komen overeen met die uit andere onderzoeken. Klinici, beleidsmakers en (toekomstige) IVF-patienten dienen zich

bewust te zijn van dit verhoogde risico bij het maken van beslissingen. Er is gesuggereerd dat het terugplaatsen van minder embryo's niet alleen het risico op meerlingzwangerschappen verkleint, maar ook het verloop van eenlingzwangerschappen gunstig zou beinvloeden (Keirse 1995). Inderdaad zijn er onderzoeksgegevens die suggereren dat vroeggeboorte en laag geboortegewicht minder frequent voor zouden kunnen komen bij eenlingzwangerschappen die tot stand komen na terugplaatsing van minder embryo's dan bij eenlingzwangerschappen na terugplaatsing van een groter aantal (Doyle 1992). Dit kan gerelateerd zijn aan de kwaliteit van de embryo's. Immers, wanneer een groter aantal wordt teruggeplaatst kan dit betekenen dat de kwaliteit van de embryo's als slecht was beoordeeld en daarmee de kans op succes per embryo als laag. Aan de andere kant zou het ook te maken kunnen hebben met factoren in de vroege ontwikkeling die ervoor zorgen dat een embryo minder optimaal groeit in het gezelschap van meerdere andere embryo's. In ieder geval zou het de moeite waard zijn om te bestuderen of bij een bewust beleid van terugplaatsing van maar een embryo, de resultaten in eenlingzwangerschappen beter zijn dan bij het huidige beleid.

De risicoverhoging die wij hebben vastgesteld bij de IVF-eenlingzwangerschappen, kan niet alleen worden verklaard uit kenmerken van de moeders zoals leeftijd en pariteit. We hebben in ons onderzoek voor een groter aantal factoren kunne corrigeren dan de meeste andere auteurs tot nu toe. Zelf na correctie voor deze factoren bleef een statistisch significant verhoogd risico op vroeggeboorte (OR=1.6 voor de vergelijking met SMOCK en OR=1.2 voor de vergelijking met de LVR) en op groeivertraging (OR=1.4) bestaan. Restrictie van onze analyses tot kleine subgroep van IVF- , LVR- en SMOCK-cohort met een laag risico op alle gemeten factoren, veranderde deze conclusie niet. Onze resultaten komen overeen met die uit twee recente onderzoeken. Bergh et al (1999) onderzochten een grote groep IVF-zwangerschappen en vonden een verhoging van het risico op vroeggeboorte en laag geboortegewicht vergelijkbaar met onze uitkomsten. In hun analyse werd getratificeerd naar meelingzwangerschap, pariteit, leeftijd, jaar van geboorte en duur van de infertiliteit. Koudstaal et al (2000a) vinden vergelijkbare uitkomsten met die in ons onderzoek in een cohort van 307 IVF-zwangeren die paarsgewijs zijn gematched met een controlegroep van niet-IVF-zwangeren op leeftijd, lengte, BMI (Body Mass Index), roken, alcoholgebruik, etniciteit and medische en obstetrische voorgeschiedenis.
Ook bij de IVF-tweelingzwangerschappen in ons bestand worden resultaten gevonden die vergelijkbaar zijn met die bij de eenlingen. IVF-tweelingen hebben een respectievelijk 1,4 en 1,3 keer verhoogd risico op laag geboortegewicht en vroeggeboorte, na correctie voor zygositeit en een aantal andere covariaten (ongepubliceerde data). Koudstaal et al (2000b) vergeleken 96 IVF-tweelingen met eenzelfde aantal controletweelingen gematched op zygositeit and een groot aantal maternale en zwangerschapsfactoren. Ook zij tonen aan dat IVF-tweelingen vaker te vroeg worden geboren en een lager geboortegewicht hebben dan controletweelingen.

Het is mogelijk dat onbekende aspecten van de onvruchtbaarheidsgeschiedenis van de IVF-moeders of van hun profiel het verhoogde risico bij IVF-eenlingzwangerschappen verklaren. De alternatieve hypothese is dat iets in de IVF-procedure het risico op negatieve zwangerschapsuitkomsten verhoogt. Als inderdaad de IVF-procedure op

zichzelf leidt tot een hoger risico, dan kan er op dit moment alleen nog maar worden gespeculeerd over welk deel van de procedure hiervoor verantwoordelijk is.
Ons onderzoek wijst erop dat een zwangerschap na cryopreservering tot betere perinatale resultaten leidt dan een reguliere IVF-zwangerschap. Dit zou erop kunnen wijzen dat de hyperstimulatie in de IVF-procedure een negatief effect heeft op de gezondheid van het embryo en het verloop van de zwangerschap. Een aanverwante verklaring zou kunnen zijn dat het baarmoederslijmvlies in ongestimuleerde cycli van betere kwaliteit is dan in gestimuleerde cycli, wat een positief effect zou kunnen hebben op de groei van het embryo. De meeste IVF-zwangerschappen na cryopreservering in ons onderzoek waren het resultaat van ongestimuleerde cycli, terwijl slechts in 4% van de reguliere IVF-cycli geen hyperstimulatie was toegepast. Hyperstimulatie kan de circulerende relaxinespiegels verhogen (Johnson 1991), terwijl hogere serum-relaxinespiegels op hun beurt weer gecorreleerd lijken te zijn met premature geboorte (Weiss 1993) en met het aantal follicels in de voorafgaande cyclus (Kristianson 1996). Olivennes (1993) vond geen verschil in vroeggeboorterisico tussen een groep van 162 IVF-patienten and 263 onvruchtbare controles die met hyperstimulatie waren behandeld. Sommige studies laten zien dat ovariele hyperstimulatie tot laag geboortegewicht en groeivertraging kunnen leiden door verhoging van het IGFBP-1 (Insuline Groei Factor Bindend Proteine) (Lino 1986, Howell 1989, Wang 1991, Johnson 1993, Johnson 1995). Een ander onderzoek suggereert dat in IVF-eenlingzwangerschappen meer abnormaal gevormde placenta's voorkomen (22% versus 6% in de controlegroep) en dat abnormale aanhechting van de navelstreng vaker voorkomt (Janiaux 1990). Bij gemanipuleerde voortplanting bij koeien en schapen leiden zowel opwekking van de ovulatie als het opkweken van het embryo tot nageslacht met een groter lichaamsgewicht, wat bij deze diersoorten een teken van pathologie is, alsmede tot meer aangeboren afwijkingen en hogere perinatale mortaliteit (Willadsen 1991, Walker 1992, Kruip 1997). De resultaten zijn slechter voor een opgewekte dan voor een natuurlijke ovulatie, en nog slechter voor een opgewekte ovulatie gevolgd door IVF (Wagtendonk 2000, Willadsen 1991). Bij gemanipuleerde voortplanting bij dieren lijdt noch de donor van de gameten noch de recipiënt aan onvruchtbaarheid. Daarom moeten deze negatieve effecten, als ze echt bestaan, worden toegeschreven aan de IVF-procedure op zich. Het is duidelijk dat er meer onderzoek nodig is naar de verschillende aspecten van de IVF-procedure en hun mogelijke negatieve effecten voor specifieke aspecten van de procedure of het risicoprofiel van de moeders als oorzaak kunnen worden aangewezen.

Er zou samenwerking moeten worden gezocht met groepen wetenschappers die onderzoek bij dieren doen omdat de omstandigheden waaronder IVF bij dieren wordt toegepast veel eenvoudiger te beheersen en te bestuderen zijn. Dit soort samenwerking zou ook licht kunnen werpen op de vraag of IVF bij de mens al of niet een verhoogd risico op geboorteafwijkingen met zich meebrengt. Als IVF-onderzoek bij dieren duidelijk maakt welke geboorteafwijkingen het meest voorkomen, zou moeten worden getracht gegevens over met name deze geboorteafwijkingen bij de mens na IVF uit een groot aantal landen samen te voegen. Alleen dan zullen de aantallen groot genoeg zijn om conclusies te rechtvaardigen over het mogelijk vaker voorkomen van afwijkingen bij kinderen die bij koeien en schapen worden gesignaleerd.

Uit ons onderzoek bleek de ontwikkeling van IVF-kinderen op een leeftijd van 2 jaar binnen de normale grenzen te zijn, al was er een verhoogd risico op een afwijkende ontwikkeling als IVF-kinderen werden vergeleken met natuurlijk verwekte kinderen. Dit verhoogde risico lijkt te kunnen worden verklaard uit de grotere kans op vroeggeboorte na IVF. Een risico op langere termijn kan niet worden uitgesloten omdat veel problemen die verband houden met ongunstige perinatale omstandigheden pas veel later aan het daglicht treden. Er zijn thans nog onvoldoende gegevens beschikbaar om het risico op problemen op langere termijn te kunnen uitsluiten.

Ook voor vrouwen die een IVF-behandeling hebben ondergaan, zijn tal van mogelijke negatieve effecten nog onvoldoende onderzocht. Een grootschalig Nederlands onderzoek is momenteel aan de gang (van Leeuwen, schriftelijke mededeling) om het mogelijke verband tussen IVF-behandeling en eierstokkanker te onderzoeken. Verder kan totdat dit grondig is onderzocht, niet worden uitgesloten dat vrouwen die een IVF-behandeling en hyperstimulatie hebben ondergaan een verhoogde kans op andere vormen van kanker hebben.

Een cruciale vraag die nog niet is beantwoord is die naar de effectiviteit van IVF. Destijds is IVF geïntroduceerd als een optie voor de behandeling van paren die geen enkele kans op spontane bevruchting hadden. Het is duidelijk dat elke zwangerschap na IVF bij vrouwen met onherstelbaar beschadigde, verstopte of afwezige eileiders toe te schrijven is aan de IVF-behandeling. Momenteel wordt echter slechts een kwart tot een derde van alle IVF-behandelingen toegepast op vrouwen met een 'zuivere' afwijking aan de eileiders. Dat betekent dat de resterende paren een kans op een spontane bevruchting hebben die boven nul ligt. Helaas is weinig onderzoek verricht waarin groepen behandelde en onbehandelde paren zijn vergeleken die met elkaar overeenkwamen in alle belangrijke opzichte afgezien van de IVF-behandeling.

Tot dusver zijn er maar drie gerandomiseerde gecontroleerde studies uitgevoerd waarin IVF-behandeling is vergeleken met een wachttijd waarin andere behandelingen plaatsvonden (Jarrell 1993, Soliman 1993, Goverde 1999). De eerste twee studies, met respectievelijk 400 en 200 participanten, vergeleken IVF met 'andere behandelingen' tijdens een wachttijd van zes maanden. De onderzoekers konden de effectiviteit van IVF alleen aantonen voor vrouwen met een eileiderprobleem. In het onderzoek van Goverde et al werden 86 paren die alleen intra-uteriene inseminatie (IUI) kregen, vergeleken met 85 paren die IUI plus hyperstimulatie kregen en 87 paren die een IVF-behandeling ondergingen, met ten hoogste zes cycli in elke groep. Het percentage zwangerschappen per cyclus was het hoogst in de IVF-groep. Het cumulatieve zwangerschapspercentage was niet hoger in de IVF-groep dan in de andere groepen doordat in de IVF-groep meer paren van verdere behandeling afzagen voordat ze het maximale aantal van zes pogingen hadden bereikt. Er is nog geen onderzoek uitgevoerd waarin IVF-behandeling is vergeleken met geen enkele behandeling. Omdat bij paren die te maken hebben met idiopatische of mannelijke verminderde vruchtbaarheid het bevruchtingspercentage naar schatting 2% per cyclus is (Goverde 1999), valt nog te bezien of de effectiviteit van IVF wel voldoende bewezen kan worden. Observationeel onderzoek waarin een IVF-groep met en controlegroep wordt vergeleken, lijdt steeds aan een kennelijke of mogelijke

onvergelijkbaarheid van de beide groepen. Zulk onderzoek kan daardoor de noodzakelijke antwoorden niet geven. Met andere woorden, ondanks de wijdverbreide toepassing van IVF en de hoge kosten die de procedure met zich meebrengt, is er nog geen sprake van dat IVF gebaseerd is op wetenschappelijke bewijzen oftewel 'evidence' (Buitendijk 1995). De vraag of deze technologie effectiever is dan een conventionele behandeling of helemaal geen behandeling is tot op heden niet voldoende beantwoord. Gegeven het feit dat het ruim twee decennia geleden is dat IVF werd geïntroduceerd, dat vele nadelige effecten inmiddels zijn gebleken en dat andere nadelige effecten (nog) niet kunnen worden uitgesloten, is dit een vrij onthutsende conclusie.

Naar mijn mening wordt het hoog tijd dat er een goed opgezet grootschalig onderzoek wordt opgestart, waaraan verschillende instellingen en mogelijk verschillende landen deelnemen, om aan te tonen of IVF effectief is en, zo ja, voor welke indicaties. Als IVF effectief blijkt te zijn voor alle indicaties of voor een subgroep, kunnen de perinatale uitkomsten van dat onderzoek worden gebruikt om mogelijke bijwerkingen te bestuderen. In een gerandomiseerd onderzoek zullen de behandelde en de onbehandelde groep veel beter vergelijkbaar zijn dan ooit het geval kan zijn in observationeel onderzoek. De groep kinderen in dit onderzoek zou nauw moeten worden gevolgd en moeten worden bestudeerd wanneer zij ouder zijn om eventuele ontwikkelingsproblemen te kunnen vaststellen, idealiter tot zij de reproductieve leeftijd hebben bereikt. De zus van Louise Brown, die eveneens met behulp van IVF werd verwekt, werd in 1999 als teenager moeder en had dus kennelijk geen vruchtbaarheidsproblemen. Een enkel geval is echter ontoereikend om te weerleggen dat de medische wetenschap met behulp van vruchtbaarheidstechnieken een nieuwe generatie van vrouwen en mannen creëert die zelf later hulp nodig kunnen hebben bij hun voortplanting. Tot dat een grootschalig onderzoek wordt opgezet zouden beleidsmakers en artsen behoedzaamheid moeten betrachten bij de toepassing van IVF.

Men zou er zelfs voor kunnen pleiten om paren alleen in aanmerking te laten komen voor een IVF-behandeling als ze ermee instellen in een studie te participeren. Deze benadering is zeker haalbaar als de wachttijd in de ene groep vergelijkbaar is met de gebruikelijke wachttijd en korter is in de andere groep. Er kan misschien tegenin worden gebracht dat het onethisch is onzekerheid te introduceren over de effectiviteit van een gevestigde procedure die hoop biedt voor zo veel paren met vruchtbaarheidsproblemen. In het onderzoek beschreven in dit proefschrift bleken onvruchtbare paren echter uitstekend in staat te zijn zich informatie over de kans op succes eigen te maken zoals die door hun specialist werd overgebracht. Zij blijken realistisch genoeg te zijn om zich bewust te zijn van de afnemende kansen met een toenemende leeftijd van de vrouw. Verder blijken jongere onvruchtbare vrouwen meer emotionele problemen te hebben en problemen met sociale interactie. De boodschap dat hun kans op spontane bevruchting (in geval van idiopathische onvruchtbaarheid) groter zou kunnen zijn dan zij dachten zou wellicht positief kunnen zijn in plaats van belastend.

Ten slotte, doorgaan met het toepassen van een behandeling die kostbaar is en emotioneel zeer belastend is voor de vrouwen die ermee te maken krijgen, terwijl er reële twijfel bestaat over de effectiviteit en vragen over mogelijke extra risico's van vroeggeboorte en

ontwikkelingsstoornissen nog niet zijn beantwoord, is onvermijdelijk de meest onethische weg. Nu IVF tot deze nieuwe eeuw is doorgedrongen, kan de tijd rijp zijn voor goed doordacht en grootschalig onderzoek waarin aandacht wordt besteed aan aspecten van effectiviteit naast economische en psychologische aspecten van de procedure alsmede medische aspecten op de korte en langere termijn. Alleen dergelijk onderzoek zal ons in staat stellen vast te stellen of IVF aan de verwachtingen beantwoordt die het 22 jaar geleden heeft opgeroepen en of IVF inderdaad onvruchtbare paren in staat kan stellen een gezond kind te krijgen op een manier die zowel ethisch als sociaal aanvaardbaar is.

Referenties

Bergh T, Ericson A, Hillensjö T, Nugren KG, Wennerholm UB. Deliveries and children born after in-vitro-fertilisation in Sweden 1982-95: a retrospective cohort study. The Lancet 1999;354:1579-85.

Buitendijk SE. Evidence-based in-vitro fertilisation. The Lancet 1995; 346: 901 (letter).

Edwards R, Steptoe P. A matter of life. 1980. Finestride Ltd. and Crownchime Ltd.

Goverde AJ, McDonnell J, Vermeiden JP, Schats R, Rutten FF, Schoemaker J. Intra-uterine insemination or in-vitro fertilisation in idiopathic subfertility and male subfertility: a randomised trial and cost-effectiveness analysis. Lancet 2000;355(9197):13-8.

Heitman E. Social and ethical aspects of in vitro fertilization. Intl J of Techn Ass in Health Care 1999; 15:1, 22-35.

Howell RJS, Economides D, Teisner B, Farkas AG, Chard T. Placental proteins 12 and 14 in pre-eclampsia. Acta Obstetricia et Gynecologica Scandinavica, 1989;68:237-40.

Jarrell JF, Labell R, Goeree R, Milner R, Collins J. In vitro fertilization and embryo transfer: A randomized controlled trial. The Online Journal of Current Clinical Trials, Document no.73, July 1993.

Jauniaux E, Englert Y, Vanesse M, Hiden M, Wilkin P. Pathologic Features of Placentas From Singleton Pregnancies Obtained by In Vitro Fertilization and Embryo Transfer. Obstetrics and Gynecology, 1990;76:61-4.

Johnson MR, Okokon E, Collins WP, Sharma V, Lightman SL. The effect of human chorionic gonadotropin and pregnancy on the circulating level of relaxin. J Clin Endocr Metabol 1991;72:1042-7.

Johnson MR, Abbas A, Norman-Taylor JQ, et al. Circulating placental protein 14 (PP14) in the first trimester of spontaneous and IVF pregnancies. Human Reproduction, 1993;8:323-6.

Johnson MR, Irvine R, Hills F, et al. Superovulation, IGFBP-1 and birthweight. European Journal of Obstetrics & Gynaecology and Reproductive Biology, 1995;59:193-5.

Keirse MJNC, Helmerhorst FM. The impact of assisted reproduction on perinatal health care. Soz Präventivmed 1995;40:343-51.

Koudstaal N, Braat DDM, Bruinse HW, Naaktgeboren N, Vermeiden JPW, Visser GHA. Obstetric outcome of singleton pregnancies after in vitro fertilization. A pair-wise matched cohort study in four Dutch University hospitals. Accepted for publication, Hum Reprod. (2000a).

Koudstaal J, Bruinse HW, Helmerhorst FM, Vermeiden JPW, Willemsen WNP, Visser GHA. Obstetric outcome of twin pregnancies after in-vitro fertilization: a matched control study in four Dutch University hospitals. Hum Reprod 2000b;15(4):935-40.

Kristiansson P, Svärdsudd, Schoultz B von, Wramsby H. Supraphysiological serum relaxin concentration during pregnancy achieved by in-vitro fertilization is strongly correlated to the number of growing follicles in the treatment cycle. Hum Reprod 1996:11:2036-40.

Kruip ThAM, Daas JHG den. In vitro produced and cloned embryos: effects on pregnancy, parturition and offspring. Theriogenology 1997;47:43-52.

Lino K, Sjoberg J, Seppala M. Elevated circulating levels of a decidual protein, placental protein 12, in preeclampsia. Obstetrics and Gynecology, 1986;68:58-60.
Olivennes F, Rufat P, André B, Pourade A, Quiros MC, Frydman R. The increased risk of complication observed in singleton pregnancies resulting from in-vitro fertilization (IVF) does not seem to be related to the IVF method itself. Hum Reprod 1993;8:1297-1300.
Soliman S, Daya S, Collins J, Jarrell J. A randomized trial of in vitro fertilization versus conventional treatment for infertility. Fertility and Sterility, 1993; 59:1239-44.
Wagtendonk van-de Leeuw AM, Mullaart E, Roos de APW, Merton JS, Daas den JHG, Kemp B, Ruigh de L. Effects of different reproduction techniques: AI, MOET or IVP, on health and welfare of bovine offspring. Theriogenology 2000; 53:575-597.
Walker SK, Heard TM, Seamark RF. In vitro culture of sheep embryos without co-culture: Sucesses and perspectives. Theriogenology, 1992; 37:111-26.
Wang HS, Lim J, English J, Irvine L, Chard T. The concentration of insulin-like growth factor (IGF-1) and insulin-like growth factor-binding protein-1 (IGFBP-1) in human umbilical cord serum at delivery: relation to fetal weight. Journal of Endocrinology, 1991;129:459-64.
Weiss G, Goldsmith LT, Sachdev R, Hagen S von, Lederer K. Elevated first-trimester serum relaxin concentrations in pregnant women following ovarian stimulation predict prematurity risk and preterm delivery. Obstet Gynecol 1993;82:821-8.
Willadsen WM, Janzen RE, McAlister RJ et al. The viability of late morulae and blastocysts produced by nuclear transplantation in cattle. Theriogenology, 1991; 35:161-170.

Afterthought

During the mid eighties I was an active member of the Dutch DES action group, which held domicile in Women's Health Centre Aletta in Utrecht. I still have a vivid memory of the discussions that took place between the women of the DES group and the women of the Centre. The Centre members' primary concern was that IVF should be available to all women, married or single, homosexual or heterosexual. Our primary concern as DES daughters was that the procedure should be safe, both for the women and the babies. The DES story had taught us that substances to which the foetus is exposed can have teratogenic effects that may not be immediately apparent at birth. Indeed in the case of DES most of the adverse effects on the unborn babies exhibited their complete impact only after the babies had reached adulthood. In our opinion at that time, the similarities between DES and some of the hormonal preparations used in the IVF treatment were large enough to warrant caution. On the one hand, the lack of research data on babies born after IVF hampered us in our efforts to substantiate our worries. On the other hand, the mere fact that no large scale, well-designed studies had been undertaken only proved in our view that caution should, indeed, be practised. In the case of DES, also, many years had passed after it's introduction before researchers first attempted to document potential longer term adverse health effects. The similarities between the DES story and the introduction of IVF seemed abundant enough that one of the women of the DES group made them the topic of her thesis (Direcks 1985). In 1984, I was invited to present a paper on the DES issue at a WHO-symposium on clinical pharmacological evaluation in drug control. In the paper I stated: "it is very often said that today we have a different attitude towards the use of drugs, especially during pregnancy, than people had 30 years ago, but I wonder whether that attitude is so completely different? Doctors today are still prescribing drugs to help their patients, often without carefully weighing all the risks and benefits. How indeed can they weigh **all** the risks and benefits if some of them may become visible only after twenty years or so?" (Buitendijk 1984).

Not long thereafter, Eylard van Hall, Professor of Gynaecology at the Leiden Academic Hospital introduced IVF as a treatment option for infertility in his clinic. Well aware of the DES story and concerned about potential health effects of the IVF procedure, he did not take this decision easily. He felt that in order to offer the complete range of treatment possibilities to the infertility patients he should introduce the procedure into his clinic. On the other hand, the potential but yet unknown side effects worried him. I had run into him on several occasions, most of them at conferences in which DES related topics were presented. He and I started discussing the necessity of well-designed follow-up studies of women and children after IVF around 1985/1986. In 1986 I was given the opportunity to be trained as an epidemiologist in the US. I left the Netherlands for four years and the plans that Professor van Hall and I had been drawing up together could, consequently, not be put into practice immediately. Upon return from the US in the summer of 1990 I was surprised to learn that although the use of IVF had taken a high flight in the Netherlands, no national follow-up study had yet been planned or started. This provided me with the opportunity, on the other hand, to contribute to such an effort. I had taken up a position at TNO Prevention and Health which appeared to be the right place from which a national study may be coordinated. Professor van Hall and I resumed our deliberations which soon resulted in a joint initiative

to call together IVF clinicians and embryologists from the licensed IVF centres in the Netherlands. The first meeting took place late 1991 in the Leiden Academic Hospital. The majority of IVF clinics were represented. Most of the representatives appeared to be in agreement that a study should be started since the discussion centred primarily around **how** and not **if** a large scale study into IVF pregnancies and babies should be performed.

A few more meetings with the IVF centres took place in 1992. During that year together with a colleague from the Institute, I wrote the research proposal. We were invited a number of times to the Ministry of Health and Welfare to discuss the funding of the proposed study. During that time, the public's concern about the potential side effects of the IVF procedure was beginning to emerge. A number of MP's asked questions in Parliament about adverse health effects possibly related to the procedure. The Secretary of State for Health felt it was the Government's responsibility to sponsor research that would be necessary to answer the questions. Research into the effects of IVF had become a political priority and in 1993 we were provided with funding to start the national study.

Almost ten years have passed since the initial preparations for this study were made. In prospective cohort studies such as this, the researcher is said to 'age together with the cohort' (Vandenbroucke 1993). The IVF children that had not yet been conceived at the time we started our study, are now five to six years old. I know I have certainly aged accordingly. In retrospect, it has all been worth it, though.

References

Buitendijk SE. DES: the time-bomb drug. In: Drugs in pregnancy and delivery, report on the thirteenth European Symposium on Clinical Pharmacological Evaluation in Drug Control. WHO;1984:53.

Direcks A. Iedereen kan toveren met een eitje; afstudeerscriptie P.V.O. Amsterdam 1985.

Vandenbroucke JP. Grondslagen der epidemiologie / J.P. Vandenbroucke, A. Hofman, W.A.J.H. van Stiphout. Uitgeverij Bunge, 1993.

Nawoord

Nawoord

Een landelijk onderzoek van deze omvang had nooit kunnen worden uitgevoerd zonder de medewerking van vele mensen. Mijn dank gaat vooral uit naar:

de IVF-moeders die met veel nauwgezetheid en enthousiasme vragenlijsten hebben ingevuld over het verloop van hun zwangerschap en de toestand van hun kinderen.

de IVF-paren die voor de start van de benadeling en anderhalf jaar later een groot aantal vragen hebben beantwoord over hun psychisch welbevinden

de tweejarige IVF- en controlekinderen die (meestal) zo goed mogelijk probeerden de door de testassistente geformuleerde opdrachten uit te voeren

de IVF-klinieken (satelliet-, transportklinieken en laboratoriumhoudende centra) die hun patienten toestemming hebben gevraagd tot deelname en gegevens over de behandelingscyclus en deels ook over de uitkomst van de zwangerschappen hebben aangeleverd

de vroedvrouwen en gynaecologen die gegevens over het verloop van IVF-zwangerschappen onder hun zorg, voor het onderzoek hebben verstrekt

de beleidsmedewerkers van het Ministerie van VWS en van het CvZ en de leden van de commissie WBMV van het CvZ die het onderzoek belangrijk genoeg vonden om te zorgen voor financiële steun en om op gezette tijden op de hoogte te worden gesteld over de uitvoering

de leden van de begeleidingscommissie (prof.dr. M. van de Bor, M.G. de Boer, mr. M. van Heurck, dr. R.A. Leerentveld, C.R. Smand, J. van der Schoor, dr. J. Vermeiden) die gedurende de looptijd van het onderzoek ons hebben laten profiteren van hun expertise op het terrein van het onderzoek

de leden van de wetenschappelijke begeleidingscommissie (prof.dr. M.J. Heineman, dr. D.D.M. Braat, dr. N. Naaktgeboren, dr. F.M. Helmerhorst, dr. P.A. van Dop) die door de Werkgroep Fertiliteit van de NVOG waren gevraagd om de gezamenlijke wetenschappelijke rapportage van de onderzoeksresultaten te begeleiden

Jan Koudstaal voor zijn klinische inzichten en zijn enthousiasme om die te incorporeren in de wetenschappelijke aanpak van het onderzoek

Paul Verkerk voor zijn belangrijke bijdrage aan het ontwerpen van het onderzoeksvoorstel, voor zijn levensreddende actie om de begroting in het format van een spreadsheet te gieten, zijn bijdrage aan de uitvoering van het onderzoek en voor zijn methodologische bijdrage aan de rapportages

Remy Hirasing voor zijn strategische inzichten wat betreft het onder de aandacht brengen van het onderzoeksvoorstel bij de beoogde subsidiegever.

Lya den Ouden en Erik Verrips voor hun inspanningen bij het ontwerpen van en publiceren over respectievelijk het tweejaar-deelonderzoek en het deelonderzoek naar gezondheidsgerelateerde kwaliteit van leven bij IVF-paren

Ton Vogels voor zijn adviezen over bestandsbeheer en het ontwerpen van data-invoerprogramma's

Josje Sterk voor de projectassistentie en de nauwgezette manier waarop zij heeft gezorgd dat alle gegevens werden verkregen, verwerkt, ingevoerd en opgeborgen en voor haar vasthoudendheid wanneer het verkrijgen van gegevens niet vanzelfsprekend bleek

Sabine Anthony voor de nauwgezette manier waarop zij alle gegevensbestanden toegankelijk heeft gemaakt voor analyses, voor het uitvoeren van ontelbare analyses, voor haar grondige administratie van deze analyses en haar bijdrage aan de wetenschappelijke rapportages

Caroline Dorrepaal en Minne Fekkes voor de uitvoering van respectievelijk het tweejaar-deelonderzoek en het onderzoek naar gezondheidsgerelateerde kwaliteit van leven, voor hun wetenschappelijk inzicht en hun bijdrage aan de analyses en rapportages

Ina Kloosterboer voor de secretariële ondersteuning in de eerste jaren van het onderzoek en Jolanda Anthonissen en Tineke de Graaf voor de secretariële ondersteuning in de latere jaren en voor alle inspanningen ten behoeve van het vervaardigen van het manuscript en het boek

Ronald Brand voor zijn heldere statistische adviezen.

Verder gaat mijn dank uit naar Jannie en Cor de Wit en naar Truus Lelyveld, die door hun liefdevolle opvang van mijn kinderen mij in de gelegenheid hebben gesteld dit onderzoek uit te voeren.

Mijn vrienden en familie wil ik bedanken voor hun steun in voor mij moeilijke tijden en voor hun vertrouwen dat het me zou lukken te promoveren.

Mijn ouders wil ik bedanken voor de vanzelfsprekendheid waarmee ze mij de gelegenheid hebben gegeven de studie te kiezen die ik wilde en voor hun geloof in mijn capaciteiten.

Curriculum Vitae

	: Geboren op 23-08-'58 te Den Haag
1977	: Eindexamen gymnasium bèta 1977, Baudartius College te Zutphen
1984	: Doctoraal examen geneeskunde, Rijksuniversiteit Utrecht
1985	: Artsexamen, Rijksuniversiteit Utrecht
1988	: Master's of Public Health, Yale University, New Haven, Verenigde Staten
1990-1997	: Senior-onderzoeker Sector Jeugd, TNO Preventie en Gezondheid, Leiden
1998- heden	: Hoofd Sector Voortplanting en Perinatologie, Divisie Jeugd, TNO Preventie en Gezondheid, Leiden