Human Reproduction vol.13 no.4 pp.783–788, 1998

DEBATE

Safety of embryo cryopreservation: Statistical facts and artefacts

Episcientific aspects of the epigenetic factors in artificial procreation

Jacques Testart

INSERM Unité 355, 32 rue des Carnets, 92140 Clamart, and Laboratory for Medically Assisted Procreation, American Hospital of Paris, 63 Boulevard Victor Hugo, 92202 Neuilly, France

This debate was previously published on Webtrack on Web1 on January 17, 1998.

For 20 years artificial methods for helping sterile couples to procreate have been expanding rapidly. In addition to hormonal treatments for induction of numerous mature oocytes at each stimulated menstrual cycle, laboratory techniques have been proposed, including principally in-vitro fertilization (IVF), which may include sperm microinjection into the oocyte (ICSI), embryo culture in medium alone or in co-culture with feeder cells, and embryo cryopreservation. These artificial procedures have to be analysed, not only with respect to their clinical efficacy (percentage of successful attempts) but also in terms of their eventual impact on various aspects of human life. In addition to potentially altering the health of patients or their offspring, these techniques may have economic, social and psychological consequences. Animal experiments are necessary prior to the clinical application of any new technique. However their results are only indicative since there are species specific factors and extrapolations to humans are more or less tentative. Moreover experimental animals are usually healthy whereas medically-assisted procreation (MAP) techniques concern individuals with abnormal performances stemming at times from genetic alterations which may interfere with offspring characteristics.

Although several studies reveal lower developmental rates for IVF and cultured mammalian embryos (Massip *et al.*, 1984), no anomalies have been reported, even with frozen–thawed embryos. The only differences found between experimental and control animals were the higher birthweight of certain newborns in bovines produced by IVF (Behboodi *et al.*, 1995) and the decreased viability of bovines or ovines in case of in-vitro procedures or nuclear transfer (Willadsen *et al.*, 1991). In fact, enhanced fetal growth can result from in-vitro culture of embryos, asynchronous embryo transfer, or progesterone treatment of the mother soon after ovulation. There is evidence to suggest that cell lineage differentiation in the manipulated embryo is altered, resulting in preferential allocation of cells to the trophectoderm and aberrant fetal growth with a larger than normal placenta. Changes in the regulation of early gene expression could result from such environmentally induced disturbance (Walker *et al.*, 1996). Nevertheless it is important to note that adult liveweights appear not to be affected and birthweight appears not to be an inheritable trait. In contrast, certain epigenetic events in the early mouse embryo can affect the adult phenotype; in nucleo-cytoplasmic hybrids, transcriptional repression of certain genes has been observed as well as mouse growth deficiency resulting in reduced adult body weight (Reik *et al.*, 1993).

The point is that these experimental results were reported at a time when MAP was already responsible for >100 000 human births. Although IVF, embryo transfer and embryo freezing–thawing procedures seem unable to induce anomalies in babies and be capable only of transmitting parental anomalies (notably the genetic anomalies which are frequent in male infertility treated by ICSI), continued investigations in animals are legitimate. This cautious approach is especially important because humans are long-lived mammals, with late puberty, and no babies have yet been born to a man or woman conceived using MAP techniques.

So, looking for normality in newborns does not suffice and recent studies demonstrating, for example, the capacity to reproduce and have normal offspring of mice born following immature sperm injection (Kimura and Yamagimachi, 1995), are welcome. Similarly, studies on the future of animals born from freeze–thaw embryos are of interest for embryologists and physicians involved in MAP. However the results of such studies are potentially disruptive in that they could seriously disturb parents who have already benefitted from MAP or those who intend to, not to mention the political and ethical authorities. For those reasons the design of such studies and the analysis of their results must leave no place for ambiguity and permit clear conclusions.

The case of embryo freezing

Several major series of frozen–thawed embryo transfers have now been carried out. The French national registry (FIVNAT, 1996) gives us the opportunity to compare >15 000 embryo transfers following freezing and nearly one hundred thousand transfers of fresh embryos. From these data only transfers of single embryos are summarized in Table I.

Significant differences were found for several parameters. Firstly the pregnancy rate was lower after transfer of a frozen embryo, as could be anticipated given that only two out three frozen embryos survive the damage associated with freezing and thawing procedures (Testart *et al.*, 1987). However, from the time of implantation, pregnancies induced with a frozen embryo were no longer at a disadvantage compared with those induced with a fresh embryo. On the contrary, frozen embryo single pregnancies were subject to fewer medical problems

Table I. Comparative results of fresh and frozen–thawed embryo transfers (from FIVNAT, 1996)

	Fresh embryos (93 015 transfers)	Frozen–thawed embryos (15 567 transfers)	Significance
Percentage pregnancies/transfers (one embryo)	9.2	7.3	**
Percentage abortions/pregnancies	18.3	20.5	NS
Percentage pathologies in single pregnancies:			
Hospitalization	12.1	7.9	**
Pre-eclampsia	4.7	3.8	
Preterm labour	9.3	5.6	**
Pathologies in singletons:			
Percentage prematurity (<37 weeks)	9.1	8.2	NS
Percentage hypotrophy (<10th percentile)	14.4	6.3	**
weight ± SD	3126 ± 560	3270 ± 516	**
Percentage malformations	2.8	2.4	NS
Percentage still births	6.3	1.6	NS
Percentage death 0–4 weeks	6.6	4.9	NS

**Significant difference (P); NS= not significant.

Table II. Various relationships in a study on the long-term effects of mouse embryo freezing (from Dulioust *et al.*, 1995)

Parameters		Influence of genotype	sex	freezing
Litter size		+	–	–
Viability		+		–
Weight:	post-natal	–	–	–
Weight:	39–67 weeks	+	+	+
Preweaning	development	+	–	+
Behaviour:	running wheel	+	+	–
	hole board	+	+	–
	open-field	+	+	–
	Krushinsky	+	+	+
	active avoidance	+	+	+
Mandible morphometry		+	+	+

+ = significant; – = not significant.

Table III. Number of studied data to assess the effects of mouse embryo freezing (from Dulioust *et al.*, 1995)

	Minimum number of studied data
Preweaning development	
2 genotypes × 2 sexes×⩾9 criteria	⩾36
Behaviour	
2 genotypes × 2 sexes × 4 periods×⩾5 tests	⩾80
Mandible morphometry	
2 genotypes × 2 sexes×⩾11 measurements	⩾44
Weight	
2 genotypes × 2 sexes×⩾3 ages	⩾12
Total number of data pairs	⩾172

and the newborns were in better condition, with higher weight at birth. This could be explained by the greater age of frozen embryo recipients and the fact that fresh embryo transfer was often the origin of a first pregnancy. From the above study and all recently published series it appears that the clinical risk of freezing human embryos is only to slightly decrease the pregnancy rate. Complementing this observation is the fact that, without cryopreservation procedures, the potential of supernumerary embryos could not be exploited.

Although no differences have been reported between children born from fresh versus frozen embryos (Olivennes *et al.*, 1996), it is not possible to ascertain whether no differences will be found in the future. In the light of a study involving mouse embryos (Dulioust *et al.*, 1995), we can anticipate epigenetic effects due to freezing and that these will be observed only in adults or in very old animals. As observed by Wood (1997) certain aspects of the experimental design of this study are a cause for concern. Notably the outbred foster mothers provided non-uniform uterine environments that could have contributed to the observed differences. Apart from this there are other possible biases which could also become increasingly common in published scientific papers.

Table II summarizes the results published by Dulioust *et al.* (1995). They observed the statistical effects of embryo freezing on mouse behaviour, mandible morphometry and weight. Surprisingly, the weight differences emerged only in very old mice. Moreover this phenomenon was not observed in females and it occurred in males of only one of the two tested strains. As shown in Table II, similar or even more dramatic differences were found by comparing groups of control mice, from non frozen embryos: the studied parameters differed depending on both mouse sex and genotype. Given these conditions it becomes less clear that cryopreservation procedures have an epigenetic effect.

Studies involving multiplication of comparisons lead inevitably to the detection of statistical differences between groups drawn from the same population. For example, given rejection of the null hypothesis at the typical 5% probability level, comparison of two experimental groups from the same homogeneous population with respect to 100 parameters will reveal differences for about five parameters.

Returning to the Dulioust *et al.* (1995) study, numerous data obtained in mice grown from frozen versus control embryos, were compared (Table III). By accumulating the parameters studied, at least 172 different data pairs were compared and this may account for the discovery of certain differences. Moreover

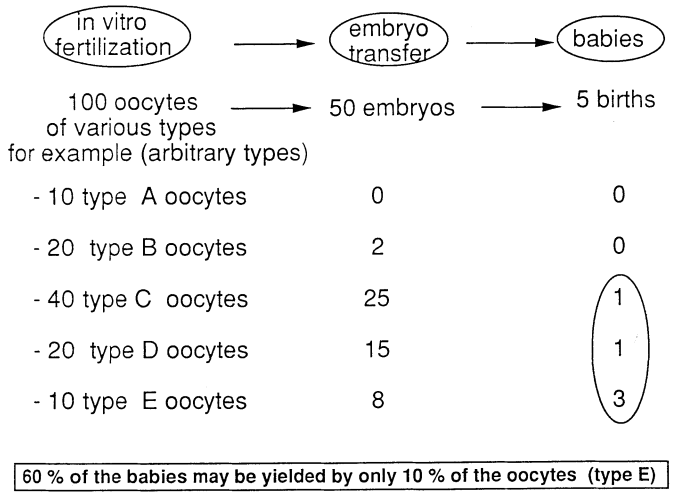

Figure 1. Hypothesis for a selection bias in artificial procreation.

it is well known that most published papers do not relate all the measures recorded, selecting only those revealing statistical differences. For example, in the paper under discussion, the mean weight of the mice was indicated in two mouse strains, for males and females, and at three different ages. One can postulate that the weight of the mice was measured at other ages, giving more chances of finding significant differences in the abundant available data.

While it is true that other studies suggest that the increased weight of in-vitro derived calves is a reproducible effect of embryo manipulations (Behboodi *et al.*, 1995; Walker *et al.*, 1996; Kruip and den Daas, 1997), they do not suffice to demonstrate that the anomaly arises from the procedure itself. The selection of embryos of different genotypes from a non-homogeneous pool may well involve a bias. Such biased selection may be innate in human artificial procreation procedures (Figure 1). From 100 recovered oocytes we currently obtain >50 embryos which give rise to about five babies. Although we are unable to discriminate between different mature oocytes, it is obvious that they have different characteristics even where they have comparable spontaneous developmental potential. It is possible that the artificial procedures, for example freezing, disfavour certain oocytes or embryos, here of types A and B (see Figure 1), but are without effect on other oocytes, namely types E and D in the example. Under such conditions, the babies born come mostly from a particular cohort of oocytes. They may have particular characteristics, such as a different mean weight, compared with babies born following natural procreation but such differences have been probably selected, rather than induced, by the artificial procreation procedures. In the Dulioust *et al.* (1995) study the effect of mouse embryo cryopreservation on preweaning development, when significant, almost exclusively concerned one of the two studied genotypes (*C3D2*). On the contrary the weight increase between 39 and 67 weeks concerned the other genotype (*B6CBA*) as though freezing–thawing procedures were able to select different individual mice according to the genotype.

In conclusion, we postulate two ways in which the effects of the diverse MAP procedures lend themselves to misinterpretation. The selective effect (Figure 1) could explain certain statistically valid, artifactual differences between experimental and control births, while the abuse of statistics and absence of confirmatory experiments may reveal false differences.The recent availability of automatic recording of experimental results, coupled with rapid, computerized data analysis in the search for statistical differences, increases the chances of finding results good enough for publication but not confirmation: the episcientific effect.

In my opinion, a replica of the Dulioust *et al.* (1995) experiment, with exactly the same procedures and measures, and even the same research team, would lead to the demonstration of other effects of freezing, different from those already reported. By playing with statistics, several of my colleagues and myself have been able to find correlations between the pregnancy rate after IVF/embryo transfer and such vitally significant parameters as the first letter of the patients names or the position of the sun or moon at the time of egg recovery. Imposition of two conditions on experimental research practises may limit misinterpretations. Firstly, rather than a politics of fishing for significant differences with a sort of statistical net, research destined for publication should involve hypothesis testing. These hypothesis may be based on rational considerations. They may also be empirical, resulting from the sort of statistical fishing practised by Dulioust *et al.* (1995). This is the second point: the production of such empirical hypotheses is a legitimate and laudable activity. However they remain to be tested and, perhaps, confirmed. Until these simple rules are respected, we can expect the epigenetic effects associated with embryo manipulation to be confounded with the episcientific effects associated with the unfortunately apt dictum: '*publish or perish*'.

References

Behboodi, E., Anderson, G., Bondurant, R. *et al.* (1995) Birth of large calves that developed from *in vitro*-derived bovine embryos. *Theriogenology,* **44,** 227–232.

Dulioust, E., Toyama, K., Busnel, M.C. *et al.* (1995) Long-term effects of embryo freezing in mice. *Proc. Natl. Acad. Sci. USA,* **92,** 589–593.

FIVNAT (1996) Bilan des transferts d'embryons congelés de 1987 à 1994. *Contracept. Fertil. Sexual.*, **24,** 700–705.

Kimura, Y. and Yanagimachi, R. (1995) Mouse oocytes injected with testicular spermatozoa or round spermatids can develop into normal offspring. *Development,* **121,** 2397–2405.

Kruip, Th. and den Daas, J.(1977). *In vitro* produced and cloned embryos: effects on pregnancy, parturition and offspring. *Theriogenology,* **47,** 43–52.

Massip, A., Van der Zwalmen, P., Puissant, F. *et al.* (1984) Effects of *in vitro* fertilization, culture, freezing and transfer on the ability of mouse embryo to implant and survive. *J. Reprod. Fertil.*, **71,** 199–204.

Olivennes, F., Schneider, Z., Remy, V. *et al.* (1996) Perinatal outcome and follow-up of 82 children aged 1–9 years and conceived from cryopreserved embryos. *Hum. Reprod.*, **11,** 1565–1568.

Reik, W., Romer, I., Barton, S. *et al.* (1993) Adult phenotype in the mouse can be affected by epigenetic events in the early embryo. *Development,* **119,** 933–942.

Testart, J., Lassalle, B., Belaisch-Allart, J. *et al.* (1987) Human embryo viability related to freezing and thawing procedures. *Am. J. Obstet. Gynecol.*, **157,** 168–171.

Walker, S., Hartwich, K. and Seamark, R. (1996) The production of unusually large offspring following embryo manipulation: concepts and challenges. *Theriogenology,* **45,** 111–120.

Willadsen, S., Janzen, R., McAlister, R. *et al.* (1991) The viability of late morulae and blastocyts produced by nuclear transplantation in cattle. *Theriogenology,* **35,** 161–170.

Wood, M.J.(1997) Embryo freezing: is it safe? *Hum. Reprod.*, **12** (Natl. Suppl.), 32–37.

No impact of cryopreservation and thawing on embryo developmental potential – one more example for the problems of retrospective, non-controlled data

Michael Ludwig[1], Safaa Al-Hasani, Ricardo Felberbaum and Klaus Diedrich

Department of Gynecology and Obstetrics, Medical University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany

[1]To whom correspondence should be addressed

This debate was previously published on Webtrack on Web2 on January 27, 1998.

In his debate article Testart (1998) claims that the statistical significance between pregnancy rates in transfer cycles using frozen–thawed embryos and fresh embryos does not reflect the different implantation rate of those embryos. He proposed that other factors are present, which might explain the obvious difference in success rates, like the rate of high quality embryos and the number of transferable embryos.

The question is difficult to answer from the general experience in freeze–thaw cycles, because the mentioned problems will always bias those data. In our mind, there are three possible ways to find a reliable answer: (i) the German experience from the transfer of frozen–thawed oocytes at the pronuclear stage, which cannot be selected by morphology; (ii) the experience from cycles, in which all embryos are frozen to avoid serious medical problems for the patient, e.g. ovarian hyperstimulation syndrome (OHSS) or due to negative predictors of success; and (iii) prospective, randomized studies, to compare directly fresh and frozen–thawed embryo transfers.

German experience with cryopreservation of oocytes at the pronuclear stage

It is well known, that embryo quality can be assessed by morphological criteria (Staessen *et al.*, 1995), but that oocytes at the pronuclear stage (PN) can not. To date we do not know any publication, which deals with morphological criteria of PN, which are helpful to assess the implantation chance. Since in Germany, due to legal restrictions, only PN can be cryopreserved, a selection bias on implantation rates is not possible. We reported recently a pregnancy rate of 17% and 18% per transfer after in-vitro fertilization (IVF) and IVF/intracytoplasmic sperm injection (ICSI) respectively, for frozen–thawed PN (Al-Hasani *et al.*, 1996). This is obviously lower than the pregnancy rate after the transfer of fresh, in the PN stage, selected embryos (26% per transfer).

The German IVF Registry (Deutsches IVF Register, 1996), which includes 2452, 11969, and 14866 transfer cycles for frozen–thawed PN, IVF and IVF/ICSI cycles, confirms a lower pregnancy rate per transfer for frozen/thawed PN cycles (10.4%), compared with fresh embryo transfers after conventional IVF or IVF/ICSI (24.1 and 23.7% respectively). The mean number of replaced embryos was similar in fresh cycles after IVF or IVF/ICSI (2.1 and 2.4 respectively) and frozen–thawed PN cycles (2.1).

Experience from cycles with threatened OHSS

The second group of studies included those with cycles, in which all embryos have been frozen, due to threatened OHSS or other problems. In all of these studies but one (Awonuga *et al.*, 1996) the clinical pregnancy rate and live birth rate were satisfactory, and the authors were convinced, that this could be a real treatment alternative for their patients in those cases.

In the analysis of cycles at high risk for OHSS from Awonuga *et al.* (1996) the clinical pregnancy (35 versus 17%; $P < 0.03$) and the live birth (27 versus 12%; $P < 0.05$) rates in patients receiving fresh embryo transfer was significantly higher than in those who had elective cryopreservation of all embryos. However, the study was not prospectively randomized. Therefore a selection bias of patients cannot be excluded.

A pregnancy rate of 38.6% was achieved by Shaker *et al.* (1996) under similar conditions in 13 patients. In another study among 23 patients at increased risk of OHSS, 15 clinical pregnancies after transfer of two to three frozen–thawed embryos in natural cycles, with a 32.6% pregnancy and 22.7% implantation rate could be achieved by Tiitinen *et al.* (1995). In 96 patients a pregnancy rate after transfer of frozen–thawed embryos of 25.2% per transfer, with a cumulative pregnancy rate of 40.6% was reported by Pattinson *et al.* (1994). Wada *et al.* (1992) reported their experience from 78 patients, who had had their embryos frozen and underwent 125 frozen–thawed embryo replacements. An implantation rate of 11% and pregnancy rates of 19 and 29% was achieved in cycles with either hormonal replacement therapy or hormonal stimulation for the frozen–thawed transfers. Finally, Frederick *et al.* (1995) also reported on a retrospective series of 36 patients, whose embryos were all cryopreserved, due to threatened OHSS or the absence of an optimal sonographic endometrial pattern at the day of transfer. After thawing a pregnancy rate of 33.3% and a live birth rate of 28.6% per cycle was achieved. The implantation rate per embryo was 9.1%, with an average number of 4.2 embryos replaced per cycle.

All but one of these publications show that there was no harm to the embryos, with regard to the capacity to implant. However, prospective randomization of fresh transfer versus all cryopreserved embryos has been carried out in only one study (Shaker *et al.*, 1996). Since these authors reported no pregnancies in fresh transfer cycles with threatened OHSS, and a pregnancy rate of 38% after freezing–thawing, these data are not really representative of day-to-day experience.

Experience from prospective, randomized studies

Selick *et al.* (1995) designed a prospective study, in which pooled fertilizable oocytes from young oocyte donors were allocated after fertilization either to fresh or frozen–thawed embryo transfer cycles. This study was controlled for male, oocyte and endometrial factors. A total of 87 transfer cycles were included in this analysis. Implantation rate per embryo and delivery rate per transfer of 12.6 and 26.2% respectively, from fresh transfer

cycles were not significantly different when compared with a per embryo implantation rate of 8.1 % and per transfer delivery rate of 13.3% from frozen–thawed transfer cycles. The lower implantation and delivery rates in the frozen–thawed group were attributed to a statistically significant difference in the number of embryos per transfer, the mean number of cells per embryo, and the rate of high quality embryos.

Horn *et al.* (1997) report on two randomly selected groups of patients. In the first group only two PN were allowed to divide, all other PN were cryopreserved (PN group). In the second group, a selection of the two best embryos was done in the cleavage stage, all other embryos were frozen (EC group). The livebirth rate per fresh embryo transfer in the EC group (27.4%) was significantly higher than that for the PN group (11.1%). Embryo survival following thawing was similar for the PN (74.4%) and EC (77.4%) stages. Although not significant, the livebirth rate following the transfer of thawed embryos was higher in the PN group (25%) than in the EC group (10.5%). Following one fresh and two freeze–thaw embryo replacements, the observed cumulative viable pregnancy rates were comparable for patients in both the PN (40%) and EC (41.1%) groups.

Conclusions

Only the German experience, published by the Deutsches IVF Register seem to support the general idea, that cryopreservation has a severe impact on the implantation rate of embryos even after exclusion of the possibility for selection. However, these data are the result of a retrospective, multicentre database, and are not controlled for embryo quality. Additionally, the impact of different methods and the reliability of the data, have at least to be discussed. Finally, a much better indicator would be the implantation rate per embryo or PN. However, these numbers are not available from the Deutsches IVF Register (1996).

The data from cycles, in which all embryos are cryopreserved due mainly to a threatening OHSS, show satisfactory results, which are in the range of those known from fresh transfers. However, these data are not randomized, as are those from Al-Hasani *et al.* (1996) and the Deutsches IVF Register (1996).

In conjunction with the results from the cycles, in which all embryos are cryopreserved, the prospective studies are of the greatest importance in answering the question. The lower pregnancy rate in the EC group after freezing–thawing, and a similar pregnancy rate after freezing–thawing in the PN group compared with the fresh transfer in the EC group show, that in fact the results of transfers after cryopreservation are biased by excluding the best embryos (Horne *et al.*, 1997). Also Selick *et al.* (1995) reported no detrimental effect of freezing–thawing on the implantation and pregnancy rates.

However, since a loss rate of 20–40% due to freezing–thawing has to be taken into account, and fresh and freeze–thaw cycles differ statistical significantly in the number of transferable embryos, their cell number and the rate of high quality embryos (Selick *et al.*, 1995), a fresh transfer should always be favoured.

The idea of a detrimental effect of the freezing–thawing procedure on human embryos seems to be wrong. However, the numbers in the cited studies were almost always very low. Therefore, to really confirm this thesis, we have to await larger, prospective studies.

This should teach us once more, that one can rely only on well designed studies, and that retrospective analysis can only help to raise questions, and not to answer them.

References

Al-Hasani, S., Ludwig, M., Gagsteiger, F. *et al.* (1996) Comparison of cryopreservation of supernumerary pronuclear human oocytes obtained after intracytoplasmic sperm injection (ICSI) and after conventional in-vitro fertilization. *Hum. Reprod.*, **11,** 604–607.

Awonuga, A.O., Pittrof, R.J., Zaidi, J. *et al.* (1996) Elective cryopreservation of all embryos in women at risk of developing ovarian hyperstimulation syndrome may not prevent the condition but reduces the live birth rate. *J. Assist. Reprod. Genet.*, **13,** 401–406.

Deutsches IVF Register (1996) *Jabrbuch 1995*. Deutsches IVF Register. Lubeck, Germany.

Frederick, J.L., Ord, T., Kettel, L.M. *et al.* (1995) Successful pregnancy outcome after cryopreservation of all fresh embryos with subsequent transfer into an unstimulated cycle. *Fertil. Steril.*, **64,** 987–990.

Horne, G., Critchlow, J.D., Newman, M.C. *et al.* (1997) A prospective evaluation of cryopreservation strategies in a two-embryo transfer programme. *Hum. Reprod.*, **12,** 542–547.

Pattinson, H.A., Hignett, M., Dunphy, B.C. and Fleetham, J.A. (1994) Outcome of thaw embryo transfer after cryopreservation of all embryos in patients at risk of ovarian hyperstimulation syndrome. *Fertil. Steril.*, **62,** 1192–1196.

Selick, C.E., Hofmann, G.E., Albano, C. *et al.* (1995) Embryo quality and pregnancy potential of fresh compared with frozen embryos—is freezing detrimental to high quality embryos? *Hum. Reprod.*, **10,** 392–395.

Shaker, A.G., Zosmer, A., Dean, N. *et al.* (1996) Comparison of intravenous albumin and transfer of fresh embryos with cryopreservation of all embryos for subsequent transfer in prevention of ovarian hyperstimulation syndrome. *Fertil. Steril.*, **65,** 992–996.

Staessen, C., Nagy, Z.P., Liu, J. *et al.* (1995). One year's experience with elective transfer of two good quality embryos in the human in-vitro fertilization and intracytoplasmic sperm injection programmes. *Hum. Reprod.*, **10,** 3305–3312.

Testart, J. (1998) Episcientific aspects of the epigenetic factors in artificial procreation. *Hum. Reprod.*, **13**, 783–785.

Tiitinen, A., Husa, L.M., Tulppala, M. *et al.* (1995) The effect of cryopreservation in prevention of ovarian hyperstimulation syndrome. *Br. J. Obstet. Gynaecol.*, **102,** 326–329.

Wada, I., Matson, P.L., Troup, S.A. *et al.* (1992) Outcome of treatment subsequent to the elective cryopreservation of all embryos from women at risk of the ovarian hyperstimulation syndrome. *Hum. Reprod.*, **7,** 962–966.

Aspects of epigenetic factors in artificial procreation

A.L.Speirs

Reproductive Biology Unit, Royal Women's Hospital, 132 Grattan Street, Carlton, Victoria 3053, Australia

[1]To whom correspondence should be addressed

This debate was previously published on Webtrack on Web2 on January 27, 1998 as a letter

For every question there is an answer that is simple, straightforward and wrong (H.L.Menken).

Testart (1998) makes a number of points which need frequent emphasis. It is now easier than ever for non-statisticians to do vast numbers of statistical tests looking for statistical significance. The discerning reader will look behind the written

word and recognize the many unreported non-significant comparisons inherent in a study design. Dr. Testart alerts us to this with regard to the mouse study by Dulioust *et al.* (1995). To demonstrate the same effect I reported a study of the letters in the surnames of pregnant and non-pregnant in-vitro fertilization (IVF) patients (Speirs, 1991). Pregnancy was strongly linked with 'GYN positive' (having G, Y or N in the surname!) $P < 0.01$. The naive reader might not realize how readily the undertaking of many statistical tests may be disguised or the way that many comparisons eventually produces a 'significant' result. This inevitability arises from the very meaning of a statistical *P* value. Incorrectly rejecting the 'null hypothesis' because of an impressive *P* value when samples are tested is described as a type I error.

It should also be recalled that statistical significance implies that an observed difference is unlikely to have arisen by chance. It should not be taken to mean causation. Because it is plausible, Dr. Testart seems to have accepted that thawed embryos produce a lower pregnancy rate than do fresh transfers because of the freeze–thaw process. This is by no means the only explanation. Single embryos transferred fresh will often be the best embryo of several available whereas a single thawed embryo will not have been selected on this basis. The thawed embryo will often be single because it is the last one available, even if nothing like the best of those originally available.

Yet another bias could be introduced by looking at who might elect to have a single embryo transferred fresh when so often many would be available. Some will be couples who particularly want to avoid twins because they have already been successful with past IVF and are returning for one more baby. Couples with past IVF success have a somewhat better chance of further pregnancy. It is not entirely the single fresh embryo that is better than the average (sometimes final) thawed embryo but their reason for electing to have a single embryo transferred when they (often) had many to choose from.

It is, of course, not possible to know from these data the extent to which any or all of these influences produced the observed difference between fresh and thawed embryo transfers.

References

Dulioust, E., Toyama, K., Busnel, M.C. *et al.* (1995) Long-term effects of embryo freezing in mice. *Proc. Natl. Acad. Sci. USA*, **92,** 589–593.

Speirs, A.L. *et al.* (1991) When predictions don't predict. *Aust. N.Z. J. Obstet. Gynaecol.*, **31,** 346.

Testart, J. (1998) Episcientific aspects of the epigenetic factors in artificial procreation. *Hum. Reprod.*, **13,** 783–785.