Clinical significance of sperm DNA damage in assisted reproduction outcome

Luke Simon¹, Gunnar Brunborg², Michael Stevenson³, Deborah Lutton⁴, Joanne McManus⁴, and Sheena E.M. Lewis¹,*

¹Centre for Public Health, Reproductive Medicine, Institute of Clinical Science, Queens University of Belfast, Grosvenor Road, Belfast BT12 6BJ, UK ²Norwegian Institute of Public Health, PO Box 4404, Nydalen, Oslo N-0403, Norway ³Centre for Public Health, Mulhouse Building, Queens University of Belfast, Grosvenor Road, Belfast BT12 6BJ, UK ⁴Regional Fertility Centre, Royal Jubilee Maternity Service, Grosvenor Road, Belfast BT12 6BA, UK

*Correspondence address. Tel: +44-28-9063-3987; Fax: +44-28-9032-8247; E-mail: s.e.lewis@qub.ac.uk

Submitted on January 27, 2010; resubmitted on March 3, 2010; accepted on March 29, 2010

BACKGROUND: Sperm DNA damage shows great promise as a biomarker of infertility. The study aim is to determine the usefulness of DNA fragmentation (DF), including modified bases (MB), to predict assisted reproduction treatment (ART) outcomes.

METHODS: DF in 360 couples (230 IVF and 130 ICSI) was measured by the alkaline Comet assay in semen and in sperm following density gradient centrifugation (DGC) and compared with fertilization rate (FR), embryo cumulative scores (ECS₁) for the total number of embryos/treatment, embryos transferred (ECS₂), clinical pregnancy (CP) and spontaneous pregnancy loss. MB were also measured using formamidopyrimidine DNA glycosylase to convert them into strand breaks.

RESULTS: In IVF, FR and ECS decreased as DF increased in both semen and DGC sperm, and couples who failed to achieve a CP had higher DF than successful couples (+12.2% semen, P = 0.004; +9.9% DGC sperm, P = 0.010). When MB were added to existing strand breaks, total DF was markedly higher (+17.1% semen, P = 0.009 and +13.8% DGC sperm, P = 0.045). DF was not associated with FR, ECS or CP in either semen or DGC sperm following ICSI. In contrast, by including MB, there was significantly more DNA damage (+16.8% semen, P = 0.008 and +15.5% DGC sperm, P = 0.024) in the group who did not achieve CP.

CONCLUSIONS: DF can predict ART outcome for IVF. Converting MB into further DNA strand breaks increased the test sensitivity, giving negative correlations between DF and CP for ICSI as well as IVF.

Key words: Comet assay / formamidopyrimidine DNA glycosylase enzyme / modified base / sperm DNA fragmentation / threshold value

Introduction

Infertility is becoming a public health issue as birth rates continue in a sustained decline across Europe. Over the last 50 years, they have plummeted to reach an unprecedented low of 1.4 children per couple (Commission of the European Communities, 2009). In 2008, the European Parliament (2008) acknowledged for the first time that falling birth rates were a major cause of its population decline. Over mortality and migration, small family size is the major determinant of the future population number and composition in Europe (Maccheroni, 2007). Infertility affects one in six couples of childbearing age (Hull et al., 1985), and male problems are responsible for 40% of these cases (Fleming et al., 1995). One solution to the problem of reduced birth rates is to lessen the decline through the use of assisted reproduction technology (ART). Europe already performs ~60% of all ART treatments in the world (Nygren and Andersen, 2001) and in European countries between 1% and 6% (Andersen and Erb, 2006; RAND, 2006) of the births are currently aided by ART. Hence, ART has the potential to significantly influence adverse economic and demographic factors, and the European parliament has finally recognized that infertility treatment should be incorporated into the proposed population policy mix (European Parliament, 2008; Ziebe and Devroey, 2008). The European Parliament (resolution adopted by Parliament on 21 February 2008) calls on Member States to ensure the right of couples to universal access to infertility treatment. If implemented, this would be a major step forward since the majority of provision for infertility is currently in the private sector (except in Scandinavia and Belgium) with only those who can afford it having access to such services.
The next step forward is for clinicians to accept the need for, and scientists to work in partnership to devise, novel diagnostic and prognostic tests to improve the relatively modest ART success rates. Mean European ‘take-home baby’ rates still have room for improvement as they are 30.1% (Andersen et al., 2008) compared with 27.0% a decade ago (Land and Evers, 2003), although some countries are more successful than others (Van den Bergh et al., 2006). The UK national live birth rate for fresh cycles to women less than 35 years is 32.3% (Human Fertilization and Embryology Authority, 2003, 2007), although it was lower in 2000 (21.8%). If ART is to be included as a substantial part of the new population policy, there will need to be government-led and -funded demands for improvement of ART success rates. Male infertility has been long neglected and this is the area where most rapid progress could be made. However, this will force ART personnel to re-examine the assessment of male fertility potential and agree on improved prognostic sperm function tests with clinical relevance for each type of ART treatment.

Conventional semen analysis by light microscopic assessment of semen parameters (semen volume, sperm count, motility and morphology) is now recognized to be of limited value in the determination of the couples’ fertility status (reviewed by Lewis, 2007). In contrast, sperm DNA testing has been increasingly recognized as a more promising test (Aitken and de Iuliis, 2007; Evenson et al., 2007; Zini et al., 2008). Measurement of sperm DNA damage is a useful biomarker for infertility with numerous studies showing its association with longer times to conceive compared with fertile couples (Spano et al., 2000), impaired embryo cleavage (Morris et al., 2002), higher miscarriage rates (Evenson et al., 1999) and also a significantly increased risk of pregnancy loss after IVF and ICSI (Zini et al., 2008). However, the implications of sperm DNA damage are even farther reaching. As sperm have few repair mechanisms (Jansen et al., 2001; Olsen et al., 2003; Aitken and Baker, 2006) and oocytes can only repair limited amount of damage (Ahmadi and Ng, 1999; Derjick et al., 2008), the damage to sperm DNA may affect the germ line for generations (Aitken and de Iuliis, 2007). Of even more concern than its ability to reduce fertility is the knowledge that sperm with oxidative DNA damage may still retain the potential to reach the oocyte, achieve fertilization and thereby contribute to mutations during embryonic development (Fraga et al., 1991) or even to loss of the fetus. If damaged sperm DNA is incorporated into the embryonic genome, it may lead to errors in DNA replication, transcription and translation during embryogenesis, contributing to a number of human diseases (Cooke et al., 2003) in not just one but subsequent generations (reviewed by Aitken et al., 2008). In particular, sperm DNA can impact on the short- and long-term health of children born by ART. Children conceived by ART, particularly ICSI, have a higher incidence of disease than those conceived spontaneously (Basatemur and Sutcliffe, 2008). Continuing into childhood, there is a strong association between poor sperm DNA integrity and diseases ranging from childhood cancers and leukemias to autism (reviewed by Aitken and de Iuliis, 2007), especially aggravated by paternal smoking (Ji et al., 1997; Sorahan et al., 1997). A number of studies have shown major congenital malformations are present in 10% of ICSI children compared with 3% in spontaneously conceived counterparts (Lie et al., 2005; Sutcliffe and Ludwig, 2007; Katari et al., 2009; Williams and Sutcliffe, 2009; Woldringh et al., 2010), whereas other reviews suggest little difference in the health of the two groups (Ludwig et al., 2006). There is controversy surrounding the assessment and clinical value of DNA assessments; however, despite the current conflict in the literature (Barratt et al., 2010; Sakkas and Alvarez, 2010), studies are rapidly accumulating (reviewed by Aitken et al., 2008) to show that the link is through DNA damage to the father’s sperm and that DNA damage is higher in ICSI patients (Bungum et al., 2007). Although there is much evidence associated with sperm DNA damage and poor ART outcomes, the tests have not been brought into clinical use.

Clinical thresholds to predict the chance of sperm populations achieving a clinical pregnancy (CP) have been established for the sperm chromatin structure assay (SCSA) (Bungum et al., 2007, reviewed by Evenson et al., 1999). A number of recent studies also show inverse relationships between fertility outcomes and DNA fragmentation (DF) using the terminal deoxynucleotidyl-transferase-mediated dUTP nick-end labelling assay (TUNEL; Spano et al., 2000; Henkel et al., 2004; Tesarik et al., 2004). As yet, there are no clinical thresholds for the Comet assay (Lewis et al., 2004), although it is recognized to be more sensitive than other DNA damage tests (Leroy et al., 1996; Irene et al., 2000) and is the only technique that allows the measurement of DNA damage in individual cells; particularly useful in a heterogeneous population such as sperm. The Comet assay measures both single- and double-strand DNA breaks using an alkaline pH method (Hughes et al., 1996; Donnelly et al., 2001). The Comet assay is highly reproducible (Hughes et al., 1997) and as it requires a much smaller number of cells (Hughes et al., 1996) for analysis than other tests, it is suitable for measures of testicular and oligozoospermic sperm samples where cells are scarce.

Oxidative stress (OS) has long been implicated as the major aetiological factor in sperm DNA damage. A low physiological level of reactive oxygen species (ROS) is accepted as necessary to maintain normal sperm function (Agarwal et al., 2003) but as ROS levels exceed physiological norms they lead to deteriorating function or reduced survival (Aitken and Baker, 2002). In contrast to somatic cells, sperm are very vulnerable to OS (Sies et al., 1992; Sies, 1993) owing to their unique membrane structures combined with limited antioxidants (Lewis et al., 1995) or protective enzymes. Not only does OS cause strand breaks but it also instigates deoxyribonucleic acid (DNA) damage to reduce DNA damage to the father's sperm.
which can be measured by the Comet assay (Collins, 2004). In order to determine both actual and potential DNA damage, we used Comet ± FPG and assessed its usefulness as a prognostic test.

Materials and Methods

Subjects

Men attending the Regional Fertility Centre, Royal Jubilee Maternity Service, Belfast, for infertility treatment between March 2008 and September 2009 were invited to participate in this study [n = 230 from IVF, mean (± SD) age 37.2 ± 0.3 years and n = 130 from ICSI, mean age 37.0 ± 0.5 years]. All subjects gave written informed consent for participation in this study, and the project was approved by the Office for Research Ethics Committees in Northern Ireland and the Royal Group Hospitals Trust Clinical Governance Committee. Semen samples were obtained after a recommended 2–5 days of sexual abstinence. All samples were subjected to a conventional light microscopic semen analysis to determine liquefaction, semen volume, sperm concentration, total sperm output and motility according to World Health Organization (WHO) recommendations (WHO, 1999). Sperm morphology was assessed according to WHO (1992) criteria. Semen analysis was performed within 1 h of ejaculation, following a period of incubation at 37°C to allow for liquefaction. After liquefaction, routine semen analyses were performed and subsequently semen was purified by density gradient centrifugation (DGC) using a two-step discontinuous Puresperm gradient (90–45%; Hunter Scientific Limited, UK). For each semen sample with a normozoospermic profile, the whole sample was layered on the top of 2 ml (90%) and 4 ml (45%) gradient and centrifuged at 250g for 20 min. For semen samples with less than normal WHO parameters, 1 ml of semen was layered on the top of 1 ml (90%) and 1 ml (45%) gradient and centrifuged at 100g for 20 min. The resulting sperm pellets were washed twice with Vitrolife GS culture media (Vitrolife Inc., Goteborg, Sweden) and concentrated by centrifugation at 250g (normozoospermic) and 100g (subnormal) for 10 min and resuspended in fresh culture media (2 ml). Hence, two populations of sperm for each patient were used to measure DNA damage by the Comet assay that with the best fertilizing potential as used for their clinical treatments (DGC sperm) and the whole population (native semen).

ART procedures

All IVF cycles were performed according to the routine procedures (Donnelly et al., 1998). Briefly, ovulation induction was achieved with recombinant FSH following a long protocol of pituitary desensitization with a GnRH analogue. HCG was administered when there were at least four follicles of diameter >17 mm, 36 h before oocyte retrieval. Mature, metaphase II oocytes obtained by vaginal ultrasound-guided aspiration were cultured in media (Vitrolife GS sequential culture media [Vitrolife Inc.]) at 37°C with 6% CO2 in air. The ICSI procedure has been described in detail previously (Van Steirteghem et al., 1993). In brief, a suspension of washed sperm was placed in polyvinylpyrrolidone (Vitrolife Inc.) and a free, motile sperm immobilized. The sperm was aspirated into the injection pipette tail-first and injected into an oocyte. Fertilization was recorded 12–16 h after injection. In each case, one or two embryos were transferred into the uterine cavity after an additional 24–48 h. Luteal phase support was provided by vaginally administered progesterone. An intrauterine pregnancy with fetal heart beat was confirmed by ultrasound 5 weeks after embryo transfer.

Single-cell gel electrophoresis (Comet) assay

Nuclear DF was assessed using an alkaline single-cell gel electrophoresis (Comet) assay as modified previously by our group (Hughes et al., 1997; Donnelly et al., 1999). Our previous study has reported an intra-assay coefficient variation of 6% for this assay (Hughes et al., 1997).

FPG treatment

Of the MB, 8-OHdG is the most commonly studied biomarker and is often selected as being representative of oxidative DNA damage owing to its high specificity, potent mutagenicity and relative abundance in DNA (Floyd, 1990). We used the protein FPG, a bacterial repair enzyme isolated from Escherichia coli, which recognizes and excises 8-OHdG generated by ROS. The FPG enzyme extract was purified from E. coli ER 2566 strain harbouring the pFPG230 plasmid, as described previously (Boiteux et al., 1990; Olsen et al., 2003). The extract has been shown to possess affinities towards the various DNA base modifications known to be recognized by pure FPG (Dr S. Sauvaigo, personal communication).

The catalytic activity of FPG involves a three-step process: (i) hydrolysis of the glycosidic bond between the damaged base and the deoxyribose, (ii) incision of DNA at basic sites leaving a gap at the 3'-ends by phosphoryl groups and (iii) removal of terminal deoxyribonucleoside 5'-phosphate from 5'-terminal site to excise the damaged base, as shown by Kuznetsov et al. (1998). To analyse MB, FPG (stock concentration of 19.14 mg/ml, diluted to a final concentration of 1 μg/ml) was added to sperm to introduce breaks at sites of MB during decondensation by lithium 3,5-diodosalicylate and incubated at 37°C for 90 min. Our previous study (Hughes et al., 1997) showed an intra-assay coefficient variation of 6% for the Comet assay. The overall SEM for all IVF/ICSI samples in this study without FPG is low (~4%) and is not increased by FPG, suggesting that variation linked to the addition of FPG is of minor importance.

Data and statistical analysis

Data were analysed using the Statistical Package for the Social Sciences (SPSS 15) for Windows (SPSS Inc., Chicago, IL, USA). Demographic details of couples are given in Table I according to the treatment (IVF or ICSI) and outcome. Our primary outcome for each treatment was the effect of DNA damage (analysed by Comet ± FPG) on CP, evaluated in both native semen and DGC sperm by logistic regression. The key outcome from the model derived above is individual posterior probabilities of a positive CP. We tested the performance of our prognostic model by calculating the c-statistic, which is identical to the area under the receiver operating characteristic (ROC) curve. Essentially, all possible pairs of individuals where one is pregnant and one is not pregnant were considered. Then, the number of such pairs where the posterior probability for the pregnant couple is higher than the posterior probability for the non-pregnant couple was counted: this was defined as the c-statistic. A null performance of the model would result in a c-statistic of 0.5.

Secondary outcomes were fertilization rate (FR) and embryo cumulative score (ECS). The FR was calculated as the percentage of all fertilized oocytes for IVF, and the percentage of metaphase II oocytes with two pronuclei for ICSI. The ECS was calculated for 153 couples who had embryo transfers on Day 3, by multiplying embryo grade (A = 4, B = 3, C = 2 and D = 1) by the number of blastomeres for each embryo and where a patient had more than one embryo, a mean across embryos was calculated to obtain the total quality of all embryos generated (ECSi) or embryos transferred (ECS). Use of ECS, as opposed to number of high-quality embryos, allows for quantification of the number and quality of blastomeres, making associations more precise. Relationships between sperm DF and the FR and ECS were compared using the Spearman rank correlation test. Associations between conventional semen parameters and DF and MB were also assessed using the Spearman rank correlation test. To determine the extent of damage contributed by MB, we compared existing DNA strand breaks with total strand breaks.
### Table I  Demographic data for couples undergoing ARTs.

<table>
<thead>
<tr>
<th></th>
<th>IVF (n = 230)</th>
<th></th>
<th>ICSI (n = 130)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pregnant</td>
<td>Non-pregnant</td>
<td>CI</td>
<td>P-value</td>
</tr>
<tr>
<td>Cycles included (n)</td>
<td>39</td>
<td>180</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Female age (years)</td>
<td>34.4 ± 4.3</td>
<td>37.5 ± 5.7</td>
<td>−11.3 to 5.1</td>
<td>NS</td>
</tr>
<tr>
<td>Number of previous treatments</td>
<td>1.5 ± 0.9</td>
<td>1.4 ± 1.2</td>
<td>−0.4 to 0.5</td>
<td>NS</td>
</tr>
<tr>
<td>Oocytes retrieved</td>
<td>10.5 ± 6.4</td>
<td>8.2 ± 5.2</td>
<td>0.4–4.2</td>
<td>0.021</td>
</tr>
<tr>
<td>Oocytes fertilized (2 pronuclei)</td>
<td>6.9 ± 3.9</td>
<td>4.9 ± 4.0</td>
<td>0.7–3.5</td>
<td>0.004</td>
</tr>
<tr>
<td>Fertilization rate (%)</td>
<td>72.6 ± 20.6</td>
<td>62.5 ± 31.2</td>
<td>−0.5 to 20.7</td>
<td>0.062</td>
</tr>
<tr>
<td>Embryos transferred</td>
<td>1.9 ± 0.2</td>
<td>1.6 ± 0.7</td>
<td>0.1–0.6</td>
<td>0.013</td>
</tr>
<tr>
<td>Total embryo cumulative score (ECS&lt;sub&gt;1&lt;/sub&gt;)</td>
<td>18.1 ± 10.9</td>
<td>13.3 ± 9.3</td>
<td>0.9–7.1</td>
<td>0.012</td>
</tr>
<tr>
<td>Transferred embryo cumulative score (ECS&lt;sub&gt;2&lt;/sub&gt;)</td>
<td>43.8 ± 17.8</td>
<td>31.0 ± 12.5</td>
<td>4.5–19.9</td>
<td>0.002</td>
</tr>
<tr>
<td>Male age (years)</td>
<td>36.1 ± 4.9</td>
<td>37.4 ± 5.0</td>
<td>−3.1 to 0.4</td>
<td>NS</td>
</tr>
<tr>
<td>Semen volume (ml)</td>
<td>3.1 ± 1.4</td>
<td>3.5 ± 2.4</td>
<td>−1.2 to 0.4</td>
<td>NS</td>
</tr>
<tr>
<td>Sperm concentration (10&lt;sup&gt;6&lt;/sup&gt; ml&lt;sup&gt;−1&lt;/sup&gt;)</td>
<td>63.7 ± 36.0</td>
<td>67.4 ± 40.7</td>
<td>−17.9 to 10.5</td>
<td>NS</td>
</tr>
<tr>
<td>Total sperm output (10&lt;sup&gt;8&lt;/sup&gt;)</td>
<td>199.3 ± 143.2</td>
<td>233.6 ± 392.5</td>
<td>−183.6 to 75.1</td>
<td>NS</td>
</tr>
<tr>
<td>Motility (%)</td>
<td>56.1 ± 20.9</td>
<td>54.8 ± 17.1</td>
<td>−5.1 to 7.7</td>
<td>NS</td>
</tr>
<tr>
<td>Normal morphology (%)</td>
<td>28.5 ± 15.2</td>
<td>24.7 ± 9.4</td>
<td>−0.01 to 7.6</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Values are the mean ± SD. NS, non-significant; CI, 95% confidence interval.
To compare the prognostic ability of the different sperm DNA damage variables, we ran logistic regression models with pregnancy (yes/no) as outcome and with each of four DNA damage markers individually as explanatory variables: the markers were DF measured by the alkaline Comet assay in native semen and in populations of sperm following DGC, with and without the addition of MB. We used the 230 IVF cycles to determine thresholds of DF (not including MB) in native and DCG sperm. Predicted probabilities of a pregnancy were obtained to achieve CP with 80% power. From these, we estimated thresholds where the predicted probability of a positive pregnancy was equal to 0.1 (ED10). Odds ratios (ORs) and their 95% confidence intervals (CIs) were computed above and below the threshold values, together with the ROC and 95% CI for ROC. Separate models were constructed for IVF and ICSI treatment groups. CIs for ROC, which include the value 0.5, are statistically indistinguishable.

Results

Comparison of conventional semen profiles from couples who achieved a pregnancy compared with couples who were unsuccessful following ART

Semen samples from couples who achieved a CP were compared with those who were unsuccessful (Table I). Abstinence times did not differ between the groups. No significant differences were observed in semen volume, sperm concentration, total sperm output and percentage motility.

Correlations between DF, FR and embryo quality assessed by the alkaline Comet assay following IVF treatment

There was no decrease in FR as DNA damage of native semen increased (data not shown, \( P > 0.05 \)). In contrast, there was a decrease in FR as sperm DF increased in DGC sperm: 0–20% and 21–40% DF were associated with higher FR (69.9 ± 3.7% and 66.4 ± 4.2%, respectively) compared with an FR of 54.4 ± 6.0% when DNA damage was 61–100% in DGC sperm (\( P < 0.05 \), Fig. 1). There was also a decrease in ECS as DF increased, both in native semen and in DGC sperm. The cumulative embryo score (ECS₁) for all the embryos generated showed a significant decrease, when DNA damage was greater than 60% in the native semen. The ECS₁ was 15.5 ± 2.6% in the group where sperm DF was 0–20% reducing to 10.7 ± 1.5% in the group where DF was 61–100% (\( P = 0.020 \)) in the native semen. The ECS₂ was only 7.3 ± 2.5% where sperm DF was 61–100% in DGC sperm (\( P = 0.032 \); Fig. 2).

Sperm DF of pregnant and non-pregnant couples following IVF

Using the Comet assay, the mean percentage of sperm DF was significantly higher in sperm from non-pregnant couples (\( n = 180 \)) compared with that from pregnant couples (\( n = 39 \)) undergoing IVF in both the native semen (51.7 ± 23.6 versus 39.5 ± 17.9; \( P = 0.004 \)) and the DGC sperm for clinical use (36.8 ± 21.6 versus 26.9 ± 14.6; \( P = 0.01 \)) (Table II). Using the threshold values of 56% for the native semen and 44% for the DGC sperm, we calculated OR and CI of 4.52 (1.79–11.92) and 6.20 (1.74–26.30), respectively, for CP (Fig. 4; Tables III and IV).

Sperm DF, FR, embryo quality and pregnancies following ICSI treatment

The ECS₂ had a significantly higher score for pregnant couples (51.1 ± 2.5) than non-pregnant couples (42.2 ± 2.3; \( P = 0.049 \)).

Figure 1 Bar chart showing decrease in fertilization rate (FR) with increase in DNA damage in the sperm prepared using DGC (sperm), for patients undergoing IVF. Values are mean ± SEM, \(* P < 0.05, n = 222.\)
from partners of couples undergoing ICSI who failed to achieve a CP tended to have more DF than sperm from pregnant couples (net increase of +8.3% native semen, \( P = 0.109 \) and +6.2% DGC sperm, \( P = 0.243 \)) (Table II). There was no correlation between sperm DNA damage and FR, or ECS1 or ECS2.

**The relationship between total DNA damage and IVF and ICSI outcomes after conversion of MB to DNA strand breaks by FPG**

A significant increase in DNA damage was detected after treatment with the DNA glycosylase FPG in both native and DGC samples (\( P < 0.0001 \); Table V). The variation in damage (with FPG) ranged from 0% to 47% in the native and 0% to 45% in the DGC sperm. In the IVF patients, addition of the FPG enzyme showed a significant increase in DF in sperm from non-pregnant (\( n = 63 \)) compared with that from pregnant couples (\( n = 10 \)) in the native semen (with a net increase of +17.1%; \( P = 0.009 \)) and in the DGC sperm (a net increase of +13.8%; \( P = 0.045 \)) (Table II). Similarly, in ICSI couples, when MB were included, the DNA damage between pregnant and non-pregnant couples was markedly different (with a net increase of +16.8% native semen, \( P = 0.008 \) and +15.5% DGC sperm, \( P = 0.024 \)) in contrast to Comet without FPG, where there was no significance (Table II).

**The prognostic value of DNA damage (strand breaks plus adducts) testing**

We tested the performance of our prognostic model by calculating the area under the ROC curve (Table II). Essentially, all possible pairs of individuals where one is pregnant and one is not pregnant
### Table II: Comparison of DF between pregnant and non-pregnant couples after IVF and ICSI treatments.

<table>
<thead>
<tr>
<th></th>
<th>IVF</th>
<th></th>
<th></th>
<th></th>
<th>ICSI</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pregnant couples</td>
<td>Non-pregnant couples</td>
<td>Difference (95% CI)</td>
<td>P-value</td>
<td>ROC (95% CI)</td>
<td>Pregnant couples</td>
<td>Non-pregnant couples</td>
<td>Difference (95% CI)</td>
</tr>
<tr>
<td>n</td>
<td>39</td>
<td>180</td>
<td>—</td>
<td>—</td>
<td>34</td>
<td>82</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>DF in native semen (%)</td>
<td>39.5 ± 17.9</td>
<td>51.7 ± 23.6</td>
<td>12.2 (−15.9, −4.0)</td>
<td>0.004</td>
<td>0.648 (0.561, 0.735)</td>
<td>58.9 ± 25.7</td>
<td>67.2 ± 25.6</td>
<td>8.3 (−18.5, 1.9)</td>
</tr>
<tr>
<td>DF in DGC sperm (%)</td>
<td>26.9 ± 14.6</td>
<td>36.8 ± 21.6</td>
<td>9.9 (−17.5, −2.4)</td>
<td>0.010</td>
<td>0.629 (0.542, 0.717)</td>
<td>45.5 ± 24.5</td>
<td>51.7 ± 27.0</td>
<td>6.2 (−16.7, 4.2)</td>
</tr>
<tr>
<td>n</td>
<td>10</td>
<td>63</td>
<td>—</td>
<td>—</td>
<td>15</td>
<td>38</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>DF in native semen after FPG treatment (%)</td>
<td>54.7 ± 4.9</td>
<td>71.8 ± 19.1</td>
<td>17.1 (−29.7, −4.4)</td>
<td>0.009</td>
<td>0.776 (0.643, 0.910)</td>
<td>63.1 ± 23.6</td>
<td>79.9 ± 18.7</td>
<td>16.8 (−29.2, −4.6)</td>
</tr>
<tr>
<td>DF in DGC sperm after FPG treatment (%)</td>
<td>42.2 ± 6.5</td>
<td>56.0 ± 19.9</td>
<td>13.8 (−27.4, −0.3)</td>
<td>0.045</td>
<td>0.693 (0.524, 0.862)</td>
<td>50.0 ± 22.2</td>
<td>65.5 ± 21.7</td>
<td>15.5 (−28.9, −2.1)</td>
</tr>
</tbody>
</table>

Values are expressed as the mean ± SD. ROC, receiver operating characteristic (area under, cm²); DGC, density gradient centrifugation; FPG, formamidopyrimidine DNA glycosylase enzyme.
were considered. The analysis for IVF CP showed the area under the curve is 0.648 cm² ($P = 0.006$) for the native semen and 0.629 cm² ($P = 0.016$) for the DGC sperm from pregnant compared with non-pregnant couples, respectively. When MB were included the area under the ROC curve increased to 0.776 cm² ($P = 0.005$) for the native semen and 0.693 cm² ($P = 0.05$) for the DGC sperm. In the ICSI group, for total DNA damage (including MB), a significant difference was observed between pregnant and non-pregnant couples. The ROC analysis for ICSI when MB were included also increased the area under the ROC curve in both the native semen and the DGC sample.

**Figure 4** Scatter plot showing DNA fragmentation (DF) measured by Comet in the native and DGC sperm according to their pregnancy outcome in IVF. $n = 219$.

<table>
<thead>
<tr>
<th>Table III</th>
<th>OR on ART outcome in IVF and ICSI cycles using a cut-off value of 56% in the native semen.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IVF</td>
</tr>
<tr>
<td></td>
<td>&lt;56%</td>
</tr>
<tr>
<td>Cycles started</td>
<td>127</td>
</tr>
<tr>
<td>Biochemical pregnancies (n, % per cycle)</td>
<td>44 (34.6)</td>
</tr>
<tr>
<td>Clinical pregnancies (n, % per cycle)</td>
<td>32 (25.2)</td>
</tr>
<tr>
<td>Deliveries to date (n, % per cycle)</td>
<td>27 (21.3)</td>
</tr>
<tr>
<td>Early pregnancy loss (n, % per cycle)</td>
<td>2 (1.6)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table IV</th>
<th>OR on ART outcome in IVF and ICSI cycles using a cut-off value of 44% in the DGC sperm.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IVF</td>
</tr>
<tr>
<td></td>
<td>&lt;44%</td>
</tr>
<tr>
<td>Cycles started</td>
<td>158</td>
</tr>
<tr>
<td>Biochemical pregnancies (n, % per cycle)</td>
<td>51 (32.3)</td>
</tr>
<tr>
<td>Clinical pregnancies (n, % per cycle)</td>
<td>36 (22.8)</td>
</tr>
<tr>
<td>Deliveries to date (n, % per cycle)</td>
<td>28 (17.7)</td>
</tr>
<tr>
<td>Early pregnancy loss (n, % per cycle)</td>
<td>2 (1.3)</td>
</tr>
</tbody>
</table>
Table V Comparison of DF in the native semen and the DGC sample according to the treatments.

<table>
<thead>
<tr>
<th>ART treatment</th>
<th>n</th>
<th>Test</th>
<th>Native semen</th>
<th>DGC sperm</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVF</td>
<td>230</td>
<td>Comet</td>
<td>49.5 ± 1.6</td>
<td>35.2 ± 1.4</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>73</td>
<td>Comet + FPG</td>
<td>74.6 ± 2.9</td>
<td>59.8 ± 3.1</td>
<td>0.0001</td>
</tr>
<tr>
<td>P-value</td>
<td></td>
<td></td>
<td>&lt;0.0001b</td>
<td>&lt;0.0001b</td>
<td></td>
</tr>
<tr>
<td>ICSI</td>
<td>130</td>
<td>Comet</td>
<td>64.0 ± 2.3</td>
<td>49.1 ± 2.3</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>53</td>
<td>Comet + FPG</td>
<td>79.4 ± 2.3</td>
<td>54.1 ± 2.4</td>
<td>0.0001</td>
</tr>
<tr>
<td>P-value</td>
<td></td>
<td></td>
<td>&lt;0.0001b</td>
<td>&lt;0.0001b</td>
<td></td>
</tr>
</tbody>
</table>

Values are the mean ± SEM.

*aP-value of comparison between native semen and DGC sperm.

Table VI Comparison of cut-off values predicting CP in the native semen (56%) and the DGC sample (44%).

<table>
<thead>
<tr>
<th></th>
<th>IVF Native (%)</th>
<th>DGC (%)</th>
<th>ICSI Native (%)</th>
<th>DGC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>82.1</td>
<td>92.3</td>
<td>47.2</td>
<td>54.6</td>
</tr>
<tr>
<td>Specificity</td>
<td>49.7</td>
<td>34.6</td>
<td>68.8</td>
<td>63.4</td>
</tr>
<tr>
<td>PPV</td>
<td>26.7</td>
<td>22.8</td>
<td>40.5</td>
<td>44.4</td>
</tr>
<tr>
<td>NPV</td>
<td>92.6</td>
<td>95.5</td>
<td>74.3</td>
<td>72.2</td>
</tr>
<tr>
<td>RR</td>
<td>3.6</td>
<td>5.0</td>
<td>1.6</td>
<td>1.6</td>
</tr>
</tbody>
</table>

PPV, positive predictive value; NPV, negative predictive value; RR, relative risk.

Discussion

In this study, the predictive value of sperm DF in native and DCG sperm on IVF and ICSI outcomes was assessed in a cohort of 360 ART patients using the alkaline Comet assay. The predictive power was significantly increased by using a modified Comet assay allowing oxidized purines to be measured, by converting such base modifications into strand breaks by means of the DNA glycosylase, FPG. We noted a relationship between existing DF and ART outcomes (FR, ECS1, ECS2 and CP) in IVF cycles but not in ICSI cycles. However, when the MB were converted into strand breaks, an inverse relationship was also observed in ICSI cycles. In the IVF group, using a threshold value of 56% and 44% DF (for native semen and DGC sperm, respectively, without FPG treatment), there was a significant decrease in pregnancy rates in the high DF group.

Male infertility diagnosis is still based on the conventional semen analysis, despite its prognostic and diagnostic limitations for the infertile couple (reviewed by Tomlinson et al., 1999; Lewis, 2007). Some studies have shown relationships between DF and sperm concentration (Tomlinson et al., 2001), normal morphology and progressive motility (Larson-Cook et al., 2003) or the absence of immature sperm (Virro et al., 2004) in native semen. Our data support those of Frydman et al. (2008) and Greco et al. (2005) in showing few correlations between conventional semen parameters and DF and is in conflict with the study reported by Irvine et al. (2000) where sperm DNA damage assessed by the Comet assay was closely associated with semen quality; in particular with sperm concentration. Here, in 68% of IVF patients, semen profiles were normozoospermic according to the WHO criteria, yet almost half of those men had DF above our threshold value of 48%. In the ICSI group, 34% had normal semen parameters (these couples had previously had IVF treatment: 47% had failed fertilization and 53% had not achieved a pregnancy), although 54% had DF above 50%. Thus, no strong relationships were found between semen analysis parameters and DF. Although conventional parameters have been shown to have no correlation with ICSI outcome (Nagy et al., 1995, 1998), these are still the characteristics by which sperm are chosen clinically, yet in this study we have again shown that these are not necessarily the sperm with the best DNA. Since sperm DNA tests show more promise, it is urgent to refine these tests until they are sufficiently robust for routine clinical use.

Sperm DNA damage has been closely associated with numerous indicators of reproductive health, including FR, ECS, implantation and spontaneous miscarriage (Lewis and Aitken, 2005; Frydman et al., 2008) using several techniques to assess sperm DNA damage. Of these, the Comet assay under alkaline and neutral conditions, TUNEL assay and SCSA (reviewed by Evenson et al., 2002; Agarwal and Said, 2003) have been shown to be most robust. Each of these tests assesses different aspects of DNA damage. The SCSA is based on partial acid-induced denaturation and staining with acridine orange, and analysis of the staining pattern of each cell using flow cytometry. On the other hand, TUNEL assay is a direct method for the assessment of DF, by quantifying the incorporated dUTP at double-strand DNA breaks catalysed by terminal deoxynucleotidyl transferase (Martins et al., 2007). The alkaline Comet assay assesses double- and single-strand DNA breaks and alkali labile sites. It has been used in vitro and in vivo in a wide variety of mammalian cells (Singh et al., 1988; Tice et al., 1990; Olive et al., 1998) employing a number of
Sperm DNA damage and assisted reproduction outcome

Sperm is associated with high DNA damage (Simon, Lewis and seen here may be expected since abnormal chromatin packing in Bakos (Tesarik et al, 2002; Seli et al, 2004; Tomlinson et al, 2002; Morris et al, 2002; Tomlinson et al, 2002; Benchaib et al, 2007; Frydman et al, 1995). The transfer of good-quality embryos is a major determinant of CP rates with IVF and ICSI (Scott, 2003; Terriou et al, 2007). In our study, sperm DF above 60% was associated with poorer ECS2 and decreasing CP in IVF but not in ICSI treatments. In IVF, successful couples had a significantly higher quality of transferred embryo (ECS2) than unsuccessful couples. However, in ICSI, the ECS2 did not differ following insemination by sperm with high or low DF. The impact of sperm DF depends on the extent of damage in the sperm and ability of the egg to repair that damage (Gandini et al, 2004). It may be that since the primary reason for these couple’s infertility is defined sperm problems, their oocytes are normal and are capable of repairing sperm DNA damage (Bungum et al, 2007; Ozmen et al, 2006). The alkaline Comet assay is highly reproducible (Hughes et al, 1997) with greater sensitivity than alkaline elution or nick translation assays even without prior chromatin decondensation (Leroy et al, 1996; Irvine et al, 2000). The Comet assay can detect damage equivalent to as few as 50 single-strand breaks per cell: another of its unique and powerful features is the ability to characterize the responses of a heterogeneous population of cells by measuring DNA damage within individual cells as opposed to just one overall measure of damaged cells versus undamaged cells, as in the TUNEL. A further advantage is that, unlike the TUNEL and SCSA which detect primarily breaks in histone-associated chromatin, the Comet assay has a broader use in detecting breaks in both protamine and histone-bound chromatin equally. One drawback of the Comet is that it requires trained researchers to perform it optimally and results can vary from lab to lab. Like the SCSA, the Comet assay would benefit from the standardized protocols and instruction from researchers who have used it extensively (e.g. those in the Robaire or Lewis laboratories).

The influence of DF on FR in assisted reproduction is still controversial since paternal DNA is not believed to influence this early fertility checkpoint. In this study, a negative correlation was observed between sperm DF and FR in both native and DGC sperm. We also observed a significant decrease in FR above 60% DF in native semen. Similarly, DGC sperm showed a significant decrease in FR with increase in DF (Fig. 1). This is in agreement with a number of studies that show a marked negative correlation between DF and FR in IVF using the TUNEL assay (Sun et al, 1997; Host et al, 2000; Benchaib et al, 2003; Huang et al, 2005; Payne et al, 2005; Borini et al, 2006; Bakos et al, 2008). Our data do not confirm those of studies using the SCSA who have used it extensively (e.g. those in the Robaire or Lewis laboratories).

The adverse effects of DF seen here may be expected since abnormal chromatin packing in sperm is associated with high DNA damage (Simon, Lewis and Oliva, unpublished results) and also with a failure of sperm DNA to decondense post-fertilization (Sakkas et al, 1996; Lopes et al, 1998). The age of the female partner has long been recognized as a significant factor in a couple’s fertility. It influences pregnancy rates after vasectomy reversal (Gerrard et al, 2007), treatment of male infertility by ICSI and even the treatment of azosperma by ICSI with surgically retrieved sperm (Silber et al, 1997). However, the data here are not an age-related phenomenon, as the ages of successful and unsuccessful women in both IVF and ICSI were similar (Table I). Another possibility is that laboratory conditions for sperm during ICSI are less deleterious than those for IVF. During ICSI, the sperm spend less time in culture media (Bungum et al, 2007) before injection into the protected environment of the oocyte and may therefore have less exposure to further oxidative damage that can occur in vitro (Dalzell et al, 2003; Agarwal et al, 2006). This possibility is supported by our data, where the addition of oxidative DNA damage, as indicated by converted MB, significantly enhances the value of the test for determining pregnancy end-points. Another option is that long periods in culture media (as more often occur in IVF) may lead to imprinting defects (Gosden et al, 2003) which could in turn impact adversely on CP rates. Since our knowledge is currently so limited as to which types of sperm DNA damage are irreparable and deleterious to reproduction, we can only speculate on this issue.

In our study, couples undergoing ICSI had lower DF than those couples undergoing ICSI (26.9 ± 14.6% versus 45.5 ± 24.5%, respectively) and attained successful CP (Table II). In IVF, DF had a significant deleterious impact on CP (Larson et al, 2000; Tomlinson et al, 2001; Duran et al, 2002; Larson-Cook et al, 2003; Saleh et al, 2003; Virro et al, 2004; Frydman et al, 2008). In contrast to our study, others have found no correlation between DF and CP (Host et al, 2000; Morris et al, 2002; Tomlinson et al, 2002; Benchaib et al, 2003, 2007; Gandini et al, 2004; Huang et al, 2005; Payne et al, 2005; Boe-Hansen et al, 2006; Borini et al, 2006; Muriel et al, 2006, 2007).
Bakos et al., 2008; Lin et al., 2008). These differences in results may be arise from variations in the assay conditions and author imposed threshold values. The threshold level for TUNEL assay varies between 4% (Host et al., 2000; Huang et al., 2005), 10% (Borini et al., 2006), 15% (Benchaib et al., 2006), 20% (Benchaib et al., 2003; Seli et al., 2004) and 35% (Frydman et al., 2008) and for SCSA, 20% (Boe-Hansen et al., 2006), 27% (Larson et al., 2000; Larson-Cook et al., 2003) and 30% (Evenson et al., 1999; Virro et al., 2004; Payne et al., 2005; Zini et al., 2005), which illustrates that there are no standardized laboratory protocols for TUNEL assay. The clinical usefulness of a test is usually based on OR which in turn are based on threshold values which vary enormously (Collins et al., 2008) depending on the assay, preparation and scientific choice. The threshold value most commonly used is that drawn from a study by Evenson et al. (1999) where 165 presumed fertile couples, none with >30% DNA damage, achieved a CP and the conclusion was therefore that >30% DF was a threshold not considered compatible with fertility. However, this threshold may or may not be appropriate for couples undergoing IVF or ICSI and may differ for sperm from native and DGC populations. In this study, two threshold values have been used to calculate ORs, 56% for native sperm and 44% for the DGC sperm. These high values (relative to the SCSA and TUNEL) are related to the sensitivity of the Comet assay in that following lysis and decondensation, all double- and single-strand breaks and alkali lable sites are revealed in contrast to other assays where perhaps only peripheral DNA damage is determined. As viability testing is not included in our standard semen analysis, another reason for the high threshold for native sperm may be the inclusion of some non-viable cells.

In ICSI cycles, in contrast to IVF, we did not obtain a significant association between DF of native or DGC sperm and CP using the Comet assay without FPG (Table II), which is in agreement with many studies (Host et al., 2000; Morris et al., 2002; Benchaib et al., 2003, 2007; Bungum et al., 2004, 2007; Gandini et al., 2004; Greco et al., 2005; Huang et al., 2005; Payne et al., 2005; Zini et al., 2005; Boe-Hansen et al., 2006; Muriel et al., 2006; Lin et al., 2008). As before, the literature is divided and, in contrast to our results, there are also studies showing a significant decrease in CP with increase in DF (Larson et al., 2000; Larson-Cook et al., 2003; Saleh et al., 2003; Virro et al., 2004; Borini et al., 2006; Bakos et al., 2008; Bungum et al., 2008). However, our data support the hypothesis that ICSI is able to compensate for existing DNA strand breaks as well as inadequate conventional sperm parameters (Ozmen et al., 2007; Bungum et al., 2008).

A major cause of sperm DNA damage is OS, caused by the generation of the ROS from contaminating leucocytes, defective sperm and antioxidant depletion (Lewis et al., 1995; Garrido et al., 2004). In addition to damage caused by creating strand breaks, we measured, for the first time, additional oxidative damage by exciting MB to make them measurable by the Comet assay. When we converted oxidized purines into strand breaks in both IVF and ICSI couples (n = 126), an increase in damage of 15.9 ± 1.3% was observed in native semen and 16.7 ± 1.4% in DGC sperm. By including MB, a strong association emerged between DF and CP rates in ICSI as well as increasing the sensitivity of detection in IVF. This shows the importance of including MB in potential prognostic tests for male infertility rather than focusing on existing strand breaks alone. Earlier studies had reported that in the measurement of MB (Horak et al., 2003a), 8-OHdG (Ni et al., 1997) is an important biomarker to investigate DNA damage and human infertility. MB are also known to increase in embryos of smoking couples (Zenzes et al., 1999). Recently, Horak et al. (2007) reported that sperm MB impairs FR during ICSI. Horak et al. (2003a) showed fertile individuals and patients with male-factor infertility differed significantly with respect to the level of bulky MB. A significant negative correlation is obtained between MB (Horak et al., 2003b), 8-OHdG (Ni et al., 1997) and semen quality in patients with an impaired fertility. Horak et al. (2003b) showed the level of bulky MB in sperm is positively associated with amounts of leucocytes in semen and also higher in semen of infertile subjects. By measuring both the DNA strand breaks and the FPG sensitive sites in human sperm, we increased the prognostic value of the Comet test. Since this study shows that a significant proportion of DNA damage is specifically a result of OS, it highlights the possibility of antioxidant therapy to protect sperm DNA prior to ART treatment.

Given the rigorous sperm selection that occurs naturally and repeatedly prior to fertilization (reviewed by Oehninger, 2000), and learning from the elegant studies of Harrison (1998) that a small minority of the unselected sperm population in semen may be normal by each assessment, it is important to examine subpopulations as well as the whole sperm population of the ejaculate. There is debate as to whether DGC isolates a subpopulation of sperm with less DF; this study supports previous work by our group (Donnelly et al., 2000) and also a report from Morrell et al. (2004) that this is indeed the case. The fragmentation (both with and without conversion of MB) of post-DGC was reduced by 10–20% (Table II). However, several studies report no differences in DNA damage in native and DGC populations (Stefanato et al., 2008; Thomson et al., 2009). As isolation of superior subpopulations has been shown to give higher ART success rates, so a clinical test for this population is needed. Density centrifugation isolates sperm with not only good functional parameters (WHO, 1999) but also better quality nuclear and mitochondrial DNA (Donnelly et al., 2000). Surprisingly, the SCSA appears only to be a useful prognostic tool for native semen (Larson et al., 2000; Bungum et al., 2008). Using the Comet assay to determine sperm DNA damage extends its usefulness to DGC populations as well as increasing its sensitivity.

Two recent systematic reviews have shown that the impact of sperm DNA damage on ART outcomes decreases from intruterine insemination to IVF and is least useful in ICSI (Collins et al., 2008; Zini and Sigman, 2009), whereas in IVF, using TUNEL and SCSA assays, the OR is 1.57 (95% CI: 1.18–2.07; P < 0.05). In our study, using DNA strand breaks only, an OR of 4.52 (1.79–11.92) in the native semen and 6.20 (1.74–26.30) in the DGC sperm for CP following IVF indicates its promise as a prognostic test. Owing to the high levels of damage observed when both strand breaks and MB were measured, it was not possible to establish thresholds for this combined test. The OR for CP following ISCI is 1.97 (0.81–4.77) in the native semen and 2.08 (0.93–4.68) in the DGC sperm showing less robustness and supporting the combined OR of 1.14 from other studies reported by Collins et al. (2008) and Zini and Sigman (2009). This supports the belief that ICSI bypasses genetic, as well as functional defects, but is difficult to explain. The conclusion is even more surprising given that all ICSI studies in the current literature
Sperm DNA damage and assisted reproduction outcome

exclude the patients’ poorest samples; the ~10% that are oligoasthenoteratozoospermic as they have no sperm surplus to their clinical requirements. This creates a bias since this group has the benefit of assisted penetration through the ISCI procedure but with relatively normal semen profiles. Perhaps the successful sperm is not typical of the cohorts analysed. If it is, the wisdom of using sperm with damaged DNA is questionable even if it does result in pregnancies, given the many animal studies showing adverse effects of DNA damage on long-term health of offspring (reviewed by Aitken et al., 2008; Fernandez-Gonzalez et al., 2008). In conclusion, this study adds to the amassing wealth of literature by showing the usefulness of sperm DNA testing in diagnosis of male infertility and that DF (or potential DF) can predict ART outcome.

Funding

We gratefully acknowledge Hamilton Thorne Biosciences for funding L.S. in his doctoral studies.

References

Ahmadi A, Ng SC. Developmental capacity of damaged spermatozoa. Hum Reprod 1999;14:2279–2285.


Sperm DNA damage and assisted reproduction outcome


Ni ZY, Liu YQ, Shen HM, Chia SE, Ong CN. Does the increase of 8-hydroxydeoxyguanosine lead to poor sperm quality? Mutat Res 1997;381:77–82.


Tesarik J, Greco E, Mendoza C. Late, but not early, paternal effect on human embryo development is related to sperm DNA fragmentation. Hum Reprod 2004;19:611–615.


Zini A, Boman JM, Belzile E, Campi A. Sperm DNA damage is associated with an increased risk of pregnancy loss after IVF and ICSI: systematic review and meta-analysis. Hum Reprod 2008;23:2663–2668.