Focus on Determinants of Male Fertility

Is sperm evaluation useful in predicting human fertility?

Sheena E M Lewis

School of Medicine, Obstetrics and Gynaecology, Queen’s University Belfast, Institute of Clinical Science, Grosvenor Road, Belfast BT12 6BA, UK

Correspondence should be addressed to S E M Lewis; Email: s.e.lewis@qub.ac.uk

Abstract

Traditionally, the diagnosis of male infertility has relied upon microscopic assessment and biochemical assays to determine human semen quality. The conventional parameters given most importance have been the concentration, motility, and morphology of sperm in the ejaculate. Most laboratories also include ‘sperm suitability’ tests where the subpopulations of sperm more likely to finish the marathon journey to the oocyte are separated by density centrifugation. These tests are essential to provide the fundamental information on which clinicians base their initial diagnosis. However, none of these parameters addresses sperm function and their clinical value in predicting fertility is questionable. The advent of intracytoplasmic sperm injection (ICSI) has further reduced the significance and perceived need for sperm quality tests since ICSI requires only one sperm, not subject to classic, or indeed any, tests for the procedure to be successful. Over the past decade, a number of laboratory tests have been developed to determine properties of sperm function. These include quantitative sperm motion parameters, capacitation, basal and induced acrosome reactions, sperm–zona pellucida interactions and nuclear and mitochondrial sperm DNA but few have been adopted into routine clinical use. International collaborations should be initiated to develop clinically relevant molecular and functional tests with agreed protocols and clinical thresholds as a matter of urgency.


Introduction

Traditionally, the diagnosis of male infertility has relied on microscopic assessment and biochemical assays to determine human semen quality. The conventional parameters given most importance have been the concentration, motility, and morphology of sperm in the ejaculate. Some laboratories have added additional tests, including estimations of vitality, anti-sperm antibodies, contaminant cells, and total motile counts before and after sperm preparation for assisted conception.

In order to establish uniformity in laboratory procedures, the World Health Organization (WHO) first published a laboratory manual for the examination of human semen and semen-cervical mucus interaction in 1980. The manual also set out standards to exclude influences such as the health of patient over the previous spermatogenic cycle, length of sexual abstinence, time, and temperature from ejaculation to analysis. The manual has been regularly updated (1987, 1999) with the latest version due out in 2008 with modifications to recommended tests for sperm morphology and motility.

The male infertile patient’s semen analysis still provides the fundamental information on which clinicians base their initial diagnosis, so it is imperative that it is performed as accurately as possible. With today’s society demanding increasing accountability from all providers of health care services, it would seem a retrograde step to accept lower standards of accuracy in the provision of andrology services. There is also an obligation for andrology laboratories to be accredited to the same exacting standards as all other medical laboratories (Clinical Pathology Accreditation, Sheffield, UK) so despite recent suggestions (Jequier 2005, 2006) questioning ‘the value of a semen analysis in any clinical situation’ and doubting that its expense was justified, it is unlikely that andrology services would be permitted to reduce the accuracy of their reports to provide approximate values or to dispense with quality control such as that required by the United Kingdom National External Quality Assessment Services for Andrology. One reason for the limited usefulness of the semen analysis is that it still not performed optimally in some laboratories (Pacey 2006) and this is a further factor in support of improving rather than downgrading current practices.

In the two decades since the WHO manuals have been our core reference points, it has become apparent that a
basic semen analysis is insufficient for the determination of the fertility status of individual men. This is hardly surprising since the WHO manual was written for a larger audience than infertility specialists including the needs of laboratories studying male reproductive toxicology. Its reference values were drawn up for populations of healthy men, not as absolute minimal values for each semen parameter necessary for conception to occur (WHO 1999). The WHO editors suggest that each laboratory should determine its own reference range for each variable. Specimens should be evaluated from men who have recently achieved a pregnancy, preferably within 12 months of the couple ceasing contraception. Those working in an infertility setting will know that this recommendation is virtually impossible to implement. Most fertile men have little motivation to participate in such an inconvenient and potentially embarrassing research study.

The poor power of semen analysis in predicting future fertility was first highlighted in the mid-1980s (Glazener et al. 1987, Polansky & Lamb 1988) and more recently by a number of groups attempting to establish more useful reference values (Bonde et al. 1998, Chia et al. 1998, Zinaman et al. 2000, Haugen et al. 2006, Iwamoto et al. 2006, Nallella et al. 2006, Swan 2006). One limitation of these studies is that all have <500 participants per study, which gives them less power. Nonetheless, the consistent message from each study has been that the WHO reference values are not predictive of fertility. This is true of both older men (>45 years; Hellstrom et al. 2006) and of young men (Jorgensen et al. 2006). This same conclusion has recently been confirmed by a larger study by Guzick et al. (2001), where two semen samples were assessed for sperm concentration, motility, and morphology from 765 male partners of infertile couples (with no female fertility concentration, motility, and morphology from 765 male partners of infertile couples (with no female fertility indication) and 696 fertile couples. They reported an extensive overlap in all three parameters between fertile and infertile men.

There is little consensus too as to which parameter within a conventional semen analysis is the best of these poor predictors of fertility. Guzick et al. (2001) reported that the percentage of sperm with normal morphology had the greatest discriminatory power, based on sensitivity and specificity of classification and regression tree analysis. A study by Chia et al. (1998) is in agreement that morphology is the best predictor, although the 243 fertile men (whose wives were pregnant at the time of semen collection) had a mean sperm morphology below the WHO cut off value for normality. Conflicting conclusions have come from a recent study by Nallella et al. (2006). Comparing 56 men of proven fertility with 406 men undergoing infertility investigations and 166 men with male factor infertility, they concluded that sperm concentration and motility were the most powerful discriminators of fertility with morphology having the poorest predictive power, since 50% of the fertile men in their study had abnormal morphology. A study by Bonde et al. (1998) was in partial agreement that the probability of conception increased with increasing sperm concentration up to $4 \times 10^9$/ml but beyond this any higher concentration was not associated with an increased likelihood of pregnancy. In contrast to Nallella et al. (2006), they found sperm motility of limited value in predicting pregnancy (Table 1).

One of the reasons for the lack of power of conventional parameters and the disagreement between studies is the inherent heterogeneity of human semen. It is one of the most variable of all biological fluids. Its conventional parameters of concentration, motility, and morphology vary significantly between countries, regions, individuals, and even between consecutive samples from one individual. When attempting to develop male ‘fertility’ predictors, another factor to consider is the regional variation in semen. For example, it is well recognized that semen quality in Finland, Estonia, and Lithuania is superior to that of Norway or Denmark (Jorgensen et al. 2006). A recent study of 99 fertile Norwegian men who had recently achieved a pregnancy (Haugen et al. 2006) has reported that the fifth and tenth percentiles for both their sperm concentration and motility fell below WHO reference values. Further, ~20% of young men from the general populations of Norway and Denmark have sperm concentrations below WHO reference levels of $20 \times 10^6$ sperm/ml, with ~40% having $<40 \times 10^6$ sperm/ml which, according to a recent publication, (Jorgensen et al. 2006) may be the ‘threshold’ below which fecundity declines.

An additional example of variations in ‘fertile’ semen parameters from continent to continent is indicated by a study from Nallella et al. (2006) who reported concentrations of $69 \times 10^6$/ml (range 48.3–120.0), 72.5 ± 16.6% motile spermatozoa and 13.8 ± 1.0% normal sperm morphology from 56 fertile men. Also in support of this regional difference, another study (Hellstrom et al. 2006) presented concentrations of $52 \times 10^6$/ml (range 33.1–84.3), and 55% (range 45–65%) motile spermatozoa in their group of 45-year-old men. Further evidence of inter- and intracontinental variation comes from Fisch et al. (1996) who have shown that sperm concentrations are markedly higher in New York ($134 \times 10^6$/ml) state than in Iowa ($48 \times 10^6$/ml). In addition, they reported a range of ‘in-between’ values from other US cities in Washington State, Texas, Minnesota, and California. They also compared sperm counts around the world and concluded that they are subject to a wide range of variation in different geographic locations varying, for example, as much as threefold between Thailand and France (Table 2).

In a cross-sectional study of men in Singapore, Chia et al. (1998) reported marked biological heterogeneity of semen even within a group of men of recently proven fertility. However, these men’s samples appeared to be more resistant to damage than those of infertile men; not showing impairment by lifestyle factors such as social alcohol consumption, cigarette smoking, or recent fever.
as has been shown to reduce the quality of semen of infertile men (Robbins et al. 2005). The heterogeneity of semen is highlighted even more dramatically in the classic graph from the WHO manual (1987, 1999) of 60 consecutive samples from one man over 120 weeks where his sperm count in one sample peaked at 170 million/ml but on seven other occasions fell below the WHO cut off of 20 million/ml. These data were confirmed by Mallidis et al. (1991) in a study of 673 samples from seven men over 324 weeks and by Alvarez et al. (2003) and Aitken & Irvine (Aitken 2006) all of which showed marked interejaculate coefficients of variation for sperm count, percentage normal morphology, and percentage motility.

### Excessive expectations of semen analysis

Recently, Jequier (2005) has questioned ‘the value of a semen analysis in any clinical situation’ doubting that its expense was justified. While she maintained semen analysis is essential for the identification of infertility and to diagnose disease severity, she stressed its limitations in diagnostic or prognostic value. Her belief is underpinned by the lack of any causal information of abnormal semen values. Semen analyses are confined to visual observations of a continually variable biological product at one point in time; giving no information as to how or why deficiencies in that product occurred. Furthermore, they provide no clue to site or time when dysfunction occurred so no therapy (if we had such an entity) can be considered. While Jequier concluded that a competent semen analysis is still essential to a man’s early investigations, she emphasized that a detailed history including occupational exposures and lifestyle hazards supported by a clinical examination of the man are also imperative before a diagnosis is made. I doubt if any andrologist would argue with that opinion. In a recent postal audit of Human Fertilization and Embryology Authority (HFEA) registered clinics in the UK by our group, we attempted to ascertain the incidence of diabetes among men attending for infertility investigations. Such a pertinent fact was available from only 2 out of the UK’s 87 units. These are important issues that we, as a scientific and clinical community providing services for infertile couples, have discussed again and again over the past 25 years but alas no funding has been forthcoming for us to address them. Although beyond the scope of this review, the role of the female partner must

---

### Table 1 ‘Normal’ semen ranges for fertile and infertile men.

<table>
<thead>
<tr>
<th>Reference</th>
<th>n</th>
<th>Sperm concentration $\times 10^6$/ml</th>
<th>Motility (%)</th>
<th>Normal morphology (%)</th>
<th>Fertility status/ART treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHO (1999)</td>
<td>1407 meta analysis</td>
<td>$\geq 20$</td>
<td>$\geq 50%$</td>
<td>$&gt; 15%^a$</td>
<td>Reference value</td>
</tr>
<tr>
<td>Fisch et al. (1996)</td>
<td>430</td>
<td>$\geq 40$</td>
<td>$&lt; 50%$</td>
<td>$20.0 \pm 10.6^b$</td>
<td>Presumed fertile</td>
</tr>
<tr>
<td>Chia et al. (1998)</td>
<td>243</td>
<td>$45 \pm 2$</td>
<td>$54 \pm 16%$</td>
<td></td>
<td>Pregnant wives</td>
</tr>
<tr>
<td>Saidi et al. (1999)</td>
<td>9612 meta analysis</td>
<td>98</td>
<td>$40 \pm 16%$</td>
<td></td>
<td>Presumed fertile from New York</td>
</tr>
<tr>
<td>Saidi et al. (1999)</td>
<td>Fertile</td>
<td>71</td>
<td>$&gt; 63%$</td>
<td>$&gt; 12^a$</td>
<td>696 fertile couples with spontaneous pregnancy within 12 months and 765 infertile couples attending for IUI</td>
</tr>
<tr>
<td>Guzick et al. (2001)</td>
<td>Indeterminate</td>
<td>765</td>
<td>$13--48$</td>
<td>$32-63%$</td>
<td>Spontaneous pregnancy</td>
</tr>
<tr>
<td></td>
<td>Subfertile</td>
<td>13</td>
<td>$&lt;32%$</td>
<td>$&lt;9^a$</td>
<td>Pregnant partners</td>
</tr>
<tr>
<td>Iwamoto et al. (2006)</td>
<td>Fertile</td>
<td>53</td>
<td>$62%$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swan (2006)</td>
<td>Fertile</td>
<td>493</td>
<td>$58.7 (53.5)$</td>
<td>Missouri</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$80.8 (64.9)$</td>
<td>California</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$98.6 (81.9)$</td>
<td>Minnesota</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$102.9 (88.3)$</td>
<td>New York</td>
<td></td>
</tr>
<tr>
<td>Haugen et al. (2006)</td>
<td>Proven fertility</td>
<td>82</td>
<td>$94.0 (71.1)$</td>
<td>$53.6 (8.1)$</td>
<td>Spontaneous pregnancy within 12 months (92%)</td>
</tr>
<tr>
<td>Nallella et al. (2006)</td>
<td>Proven fertile</td>
<td>56</td>
<td>$69.9 (48.3-120.0)$</td>
<td>$72.5 \pm 16.6$</td>
<td>Spontaneous pregnancy within 24 months</td>
</tr>
<tr>
<td></td>
<td>Donor</td>
<td>91</td>
<td>$65.0 (40.7-95.7)$</td>
<td>$72.9 \pm 15.7$</td>
<td>Healthy, unproven fertility</td>
</tr>
<tr>
<td></td>
<td>Subfertile</td>
<td>166</td>
<td>$21.3 (7.9-44.2)$</td>
<td>$37.0 \pm 20.6$</td>
<td>Attending for infertility investigatations with no female factor</td>
</tr>
<tr>
<td>Hellstrom et al. (2006)</td>
<td>247</td>
<td>$52 (33-84)$</td>
<td>$55 (45-65%)$</td>
<td>$59 (53-66)^b$</td>
<td>Healthy men 45-47 years</td>
</tr>
</tbody>
</table>

Ranges or s.d. in parentheses.

$^a$By Tygerberg criteria. $^b$By WHO (1987).
be considered in our assessment of potential male fertility. The length of time the couple have been infertile should also be included as part of the diagnosis as this is known to impact on assisted reproductive technology (ART) success (Basso & Baird 2003, Basso & Olsen 2005). To date, there is no formulaic consideration to include all these crucial factors, so it is of little surprise that our tests are lacking in predictive value (Holt 2005).

Declining sperm counts

Since the controversial paper by Carlsen et al. (1992) reporting that sperm counts in the western world had decreased by 1% per year over the last 50 years, there has been a large body of additional studies supporting this belief (Ginsburg et al. 1994, Auger et al. 1995, Adamopoulos et al. 1996, Irvine et al. 1996). Interpretation of these data and their statistical significance are still questioned and disbelieved by other investigators (Younglai et al. 1998, Andolz et al. 1999, Jouannet et al. 2001, Handelsman 2001). Measures of sperm concentration alone are not necessarily the definite biomarker of testis function nor can we extrapolate from declining sperm concentration to male infertility given that the human male produces up to $500 \times 10^6$ in each ejaculate and even with a 50% reduction leaves a significant excess. Further, the sperm analyzed from either whole semen samples or the subpopulations prepared for assisted conception may bear little resemblance to the characteristics of the tiny population of sperm reaching the oocyte in the Fallopian tube (Templeton & Mortimer 1982).

A more appropriate test may be of the numbers of functionally competent sperm with most potential for fertilization (see sperm function tests). Nonetheless, male reproductive health does appear to be under threat and concomitant increases in male congenital abnormalities—hypospadias and cryptorchisms and carcinomas—in situ have led Skakkebaek's group to postulate 'the testicular dysgenesis syndrome' in which all of these deteriorating trends are associated with environmental exposure to endocrine disruptors, particularly in utero (Bay et al. 2006).

Factors affecting evaluation of sperm quality

Recent reviews have also highlighted some of the dangers of laboratory handling to gametes destined for clinical use (Lewis & Aitken 2005, Agarwal et al. 2006). Not only may these procedures affect our semen analyses, but they can also compromise sperm function and induce DNA fragmentation in the laboratory.

---

**Table 2 Regional variation in semen quality.**

<table>
<thead>
<tr>
<th>Study</th>
<th>n</th>
<th>Fertility status</th>
<th>Concentration × 10^6/ml</th>
<th>Motility (%)</th>
<th>Normal morphology (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norway</td>
<td>Haugen et al. (2006)</td>
<td>82</td>
<td>Spontaneous pregnancy within 12 months (92%)</td>
<td>94 (71)</td>
<td>54 (8)</td>
</tr>
<tr>
<td>UK</td>
<td>Irvine et al. (1996)</td>
<td>577</td>
<td>Healthy donors</td>
<td>98±78</td>
<td>66±12%</td>
</tr>
<tr>
<td>France-urban-Paris</td>
<td>Auger &amp; Jouannet (1997)</td>
<td>1396</td>
<td>Proven fertile semen donors</td>
<td>98±73</td>
<td>66±14%</td>
</tr>
<tr>
<td>France-rural-toulouse</td>
<td>Auger &amp; Jouannet (1997)</td>
<td>371</td>
<td>Proven fertile semen donors</td>
<td>85±</td>
<td>66±14%</td>
</tr>
<tr>
<td>USA-urban</td>
<td>Fisch et al. (1996)</td>
<td>1400</td>
<td>Proven fertile</td>
<td>79–134</td>
<td></td>
</tr>
<tr>
<td>USA-urban</td>
<td>Nallella et al. (2006)</td>
<td>56</td>
<td>Spontaneous pregnancy within 24 months</td>
<td>70 (48–120)</td>
<td>72±17</td>
</tr>
<tr>
<td>USA-rural</td>
<td>Fisch et al. (1996)</td>
<td>386</td>
<td>Presumed fertile</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>USA-regions</td>
<td>Swan (2006)</td>
<td>493</td>
<td>Pregnant partners</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>Iwamoto et al. (2006)</td>
<td>324</td>
<td>Partners currently pregnant without intervention</td>
<td>53</td>
<td>62%</td>
</tr>
<tr>
<td>Singapore</td>
<td>Chia et al. (1998)</td>
<td>243</td>
<td>Partners currently pregnant without intervention</td>
<td>45±3</td>
<td>55±16%</td>
</tr>
<tr>
<td>Denmark</td>
<td>Jorgensen et al. (2001)</td>
<td>1082</td>
<td>Fertile</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>Jorgensen et al. (2001)</td>
<td>1082</td>
<td>Fertile</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>Thailand</td>
<td>Fisch et al. (1996)</td>
<td>307</td>
<td>Presumed fertile</td>
<td>53</td>
<td></td>
</tr>
</tbody>
</table>

s.d. in parentheses.  
aBy Tygerberg criteria. bBy WHO (1987).
We begin our preparation by routinely removing the sperm’s antioxidant defense which, in terms of chain-breaking antioxidants, predominates in the seminal plasma rather than the sperm per se (Lewis et al. 1995). This makes those samples already generating excessive reactive oxygen species (ROS; Aitken et al. 1992) even more vulnerable to oxidative stress. Despite the clear message of the error of centrifuging semen with good and poor quality sperm together (Aitken & Clarkson 1987) nearly 20 years ago, such practices still prevail, especially for samples displaying oligoastenoteratozoospermia which should be given more protection. The atmospheric environment of the laboratory is not optimal for gamete viability particularly with pO_2, which is 20-fold higher than intracellular pO_2 (Agarwal et al. 2006). This is supported by recent work on stem cells showing that 3–8% pO_2 are optimal levels whereas routine laboratory is 20% pO_2 at 37 °C and clearly deleterious for stem cells and gametes (Sullivan et al. 2006). A further hazard to gametes is present in the laboratory in the form of light (Agarwal et al. 2006). Even visible light (400–700 nm) radiations together with molecular oxygen can produce oxidative damage to vulnerable cells (Foote 1968). Membrane lipids are photosensitized by oxidation with lipid hydroperoxides acting as intermediates of the peroxidative process that leads to irreversible damage (Girotti 2001). Sperm membranes are particularly vulnerable due to their high unsaturated lipid content.

Temperature and pH are further influences on gamete stability and potential (Hamamah & Gatti 1998; reviewed by Bedford 2004) but as yet we have not developed laboratory protocols to ensure their protection. Testicular sperm appear to be more susceptible to damage than the more mature ejaculated sperm yet they are subjected to conditions which assume that they have similar resistance to injury. For example, incubation under aerobic conditions for short (4 h) or long (24 h) at 37 °C leads to marked sperm DNA damage (Dalzell et al. 2003, Dalzell et al. 2004). Testicular sperm also manifest significant cryoinjury whether they originate in fertile or infertile men unlike ejaculated sperm from fertile men which appear to be resistant to such assault (Dalzell et al. 2004). Some cryopreservation protocols have been shown to induce irreversible damage to testicular sperm (Thompson-Cree et al. 2003) while others are safer, yet there is no consensus as to which is used clinically.

The lack of consensus on appropriate sperm tests and the clinical reluctance to use even those with proven benefit is preventing patients making informed choice as to whether they should proceed with treatment. While patients have a right to proceed even if chances of success are low, particularly as the majority of cycles are self-funded, they should be given the maximum information available to inform their choices. In one study, Liu & Baker (2004) have shown that 80% of males with oligozoospermia have sperm incapable of fertilization. Thus, such patients might be recommended to bypass in vitro fertilization (IVF), where fertilization failure is likely, as their first treatment and have intracytoplasmic sperm injection (ICSI) instead allowing them to avoid unnecessary emotional and financial burdens. However, common clinical practice, if sufficient normal sperm are available, is to proceed to ICSI only when the couple has been unsuccessful.

**Assessment of nuclear and mitochondrial DNA**

Among the tests showing most promise in predicting the successful treatment of male infertility patients are those measuring sperm DNA quality (Agarwal & Said 2004, Lewis & Aitken 2005).

Sperm DNA damage has also been shown to have the lowest variability of all semen parameters examined, exhibiting a coefficient of variation of around <10% (Schrader et al. 1988, Evenson et al. 1991), although one study has recently reported it to be 29% (Erenpreiss et al. 2006). Thus, many consider it to have potential as a more useful diagnostic tool in the clinical assessment of semen quality (Zini et al. 2001, Loft et al. 2003, Holt 2005).

In recent years, the rapid advance of molecular biology has resulted in numerous techniques to assess DNA and chromatin quality. Of these, the Comet, TUNEL, and sperm chromatin structure assays (reviewed by Agarwal & Said 2005, Lewis & Aitken 2005, Aitken & De Iulissi 2007, Evenson et al. 2007) have been shown to have the strongest prognostic power. However, the power of an individual test may vary depending on the form of ART which is employed (e.g. Bungum et al. 2007).

**Sperm function tests**

As well as possessing ‘good enough’ DNA, the sperm must be functionally competent with the intrinsic
capabilities to deliver its genome to the oocyte. The fertilization process is a net result of a series of molecular events enabling sperm to reach, recognize, bind to, and penetrate the oocyte. To this end, a number of effective sperm function tests have been developed (reviewed by Aitken 2006). Measurement of sperm motion parameters using computer-aided sperm analysis (CASA) has improved the precision and reproducibility of the values measured and facilitated quantitative determinations of velocity and characteristics of track direction. It provides rapid, objective assessments that are difficult to attain with light microscopy. A number of studies have reported strong correlations between CASA parameters; particularly straight, path and curvilinear velocities and fertilization in vitro and pregnancy (Donnelly et al. 1999, Hirano et al. 2001, Silva et al. 2006). Holt et al. (1997) performed an elegant series of studies using animal sperm where fertility endpoints can be monitored much more accurately and demonstrated strongly significant associations between sperm motility and conception ($P<0.0001$).

Hyperactivation, the end stage, non-progressive motility the sperm exhibits on arrival at the oocyte, has also been shown to correlate with its ability to fertilize the oocyte (Sukcharoen et al. 1995), although in conflict another study (Yogev et al. 2000) showed no association between hyperactivation and zona binding. Unlike progressive motility, this action is characterized by high curvilinear velocities and lateral head displacement (ALH) which enable the sperm to transverse the dense zona pellucida (ZP). However, in the laboratory, these properties make it difficult to measure accurately as it moves in and out of the field of view. Similarly, the sperm’s ability to penetrate cervical mucus or hyaluronic acid substitutes has also been flagged as a useful measure of its fertilizing potential in vitro (Eggert-Kruse et al. 1989, Sharara et al. 1995). Its relevance in vivo has yet to be ascertained.

Other properties that remain invisible (Holt 2005) unless appropriately challenged may also be of prognostic value. These include sperm–zona recognition and penetration (Liu & Baker 2004) and acrosome reactions. The basal acrosome reaction has limited usefulness (Plachot et al. 1984) but acrosome reactions induced by the ionophore A23187 (ARIC) test is a good predictor of the sperm’s fertilizing potential and is markedly reduced in infertile men (Cummins et al. 1991). The acrosome reaction to progesterone is even more strongly correlated to fertilization rates in vitro (Krausz et al. 1996) and when the two (A23147 and progesterone) are combined the positive predictive value increased to $>95%$.

It has recently been used in many studies as an important biomarker (Whan et al. 2006, Glenn et al. 2007). Acrosome-reacted sperm have also been shown to lack the ability to bind to the ZP (Liu et al. 2004) so are essentially infertile sperm. In one study (Liu et al. 2001) of 186 men with normozoospermia but unexplained infertility, 54 of them had disordered ZP-induced acrosome reactions. The authors suggest that, for such men, this test should be performed as these patients will be always unsuccessful with IVF and should be referred directly for ICSI (Liu et al. 2001). However, there is a wide range of ZP-induced acrosome responses (20–98%; Liu et al. 2003) even within fertile normozoospermic men which makes it difficult to obtain clinically relevant threshold values.

**Oxidative stress tests**

Of all the tests discussed so far, none have been shown to be both sufficiently strong markers or reproducibly to be adopted into clinical practice. This had led to new avenues of investigation for more useful tests. In particular the assessment of ROS that are known to be necessary for the normal functioning but at elevated levels also a common cause of malfunctioning of sperm. Aitken et al. (2003) have shown that sperm have multiple plasma membrane redox systems involved in physiological control of the cell. Another ROS, nitric oxide (NO) has been shown to be necessary for sperm motility but is deleterious; particularly in human sperm, unless maintained within a tightly controlled physiological concentration range (Donnelly et al. 1997).

NO synthase (NOS) has been found in the heads and midpiece regions of sperm with more intense immuno-fluorescence of eNOS and bNOS in normozoospermic samples (Lewis et al. 1997) supporting a physiological role for NO.

However, oxidative stress from excess ROS, such as hydrogen peroxide, superoxide anions, and hydroxyl radicals present either from increased production or reduced antioxidant protection (Aitken et al. 1992, Lewis et al. 1995) are thought to be a major cause of sperm dysfunction. One of the primary mechanisms by which this occurs is via the stimulation of a lipid peroxidation cascade in the plasma membrane (Aitken & Clarkson 1987, Aitken et al. 1989). Sperm are particularly susceptible to ROS-mediated injury due to their high level of unsaturated fatty acids and inability to repair damage. Numerous studies have reported associations between oxidative stress and structural (Cummins et al. 1994, Twigg et al. 1998, Irvine et al. 2000) and function damage (reviewed by Aitken & Fisher 1994). Tests to determine potential and executed oxidative injury are chemiluminescence tests to assess ROS generation (Donnelly et al. 1999) lipid peroxidation tests to determine the irreversible damage-induced (Agarwal et al. 2006) and antioxidant capacity tests (Lewis et al. 1995) to measure the sperm and seminal plasma protection against ROS. The strong prognostic value of these tests on ART outcomes has been reported by a number of studies (Aitken & Baker 2006).
Urgency of appropriate test implementation

One of the reasons for the delay in implementing appropriate sperm function tests has been the success of ICSI for men with male infertility. The studies of Nagy et al. (1995) and Svalander et al. (1996) have had enormous influence in clinics around the world in showing that immediate outcome (in terms of implantation – not child health) did not relate to conventional semen analysis and from then semen analysis of poor samples destined for ICSI was performed in a cursory way only to estimate the approximate number and quality of sperm present in the ejaculate. Since ‘take-home baby’ rates are similar for IVF where the best sperm are isolated and presented to the oocyte as for ICSI where any sperm, even those which would not normally have the capacity to fertilize, are injected into oocytes bypassing all selection barriers, there is little incentive for infertility centers to invest in the research and development of sperm selection tests. Although this policy is prevailing in the short term, we have little knowledge as to the effects of this action on the health of future generations. We do know that in those cases where spermatogenic compromise is due to Y chromosomal microdeletions in the father, these genetic defects are likely to cause male infertility in their sons (Oates et al. 2002). Although ICSI has provided treatment for new groups of couples with male infertility who were previously untreatable by IVF it has not led to a marked overall improvement in assisted conception success rates. In order to treat couples with the least invasive treatment, better prognostic tests for fertile sperm, oocytes and embryos are needed. The European average ‘take-home baby’ rates remain similar now (ESHRE 2002) to those a decade ago (Van Steirteghem et al. 1993, Harari et al. 1995), although some countries have reported higher pregnancy rates (Van den Bergh et al. 2006). In UK, the national live birth rate for fresh cycles to women under 35 years is still 27% (HFEA 2005–2006).

This acceptance of static success rates may change as a result of economic pressures across Europe to increase birth rates. ‘Trends in pregnancy and fertility rates’ was the theme of a recent meeting (the Third International Workshop of Environment, Reproductive Health and Fertility, Copenhagen 2006) where birth rates were reported to be declining at an unprecedented rate so that instead of 2.1 children/couple as is necessary to maintain replacement (Lutz 2006) current rates stand at 1.4 children/couple (EuroStat, European Union Statistical Commission, 2006; RAND 2006, http://www.rand.org). This trend is multifactorial. Changes in women’s roles in society and the choice of couples to be childless have undoubtedly contributed to this reduced fecundity. It has also been impacted by individual countries’ health policies of contraception, sterilization, and abortion. However, as these social trends have not altered significantly (Jensen et al. 2002), it is more likely that falling birth rates are due to an increase in infertility (Skakkebaek et al. 2006). With predictions of European senior citizens increasing by 60% by 2050 while the working population will reduce by 18% many strategies to increase birth rates are being considered. These include a range of fertility-friendly public policies such as low mortgages on family homes and improved maternity benefits (reviewed by Lutz 2006; RAND 2006, http://www.rand.org). Short-term measures to increase the western European work force already include replacement immigration from eastern European countries but since this is contrary to Europe’s current socio-political agenda, novel measures are needed instead. One of these is clearly an increase in birth rates through ART. This is clearly a long-term policy since the impact of ART births will not influence the work force for the next decade. Nonetheless, Europe performs about 60% of all ART treatments in the world (Prof. Karl Nygren and Dr Anders Nyboe Andersen, the fourth ESHRE report on ART in Europe) and between 1 and 6% (Andersen & Erb 2006, RAND 2006, http://www.rand.org) of European births are already aided by ART. Further government-funded treatment could encourage childbearing and add a small but significant number of births to alleviate the problem. This approach is already being undertaken in Estonia (RAND 2006, http://www.rand.org) and in Belgium (Ombelet et al. 2005), where in July 2003 the Belgian government began a program of reimbursement of a maximum of six free IVF/ICSI cycles/couple. A major disincentive to governments funding ART is the substantial perinatal cost of multiple pregnancies (Callahan et al. 1994). In the ‘Belgian project’, to prevent this, only single embryo transfers were included.

If these new initiatives are adopted by other countries in Europe, there will soon be government-led demands for improvement of ART success rates. This will require us to revisit assessment of male fertility potential and agree sperm structure and function tests with established prognostic tests with thresholds of clinical relevance for each ART intervention. These should be compared with the time taken for apparently fertile couples to conceive. When the ‘minimum set of tests with maximum functional coverage to establish a practical and cost-effective service’ (Holt 2005) has been agreed, it should be followed with multi-centre trials using standardized protocols and data interpretation, training and monitoring. Such endeavors are long overdue.

Acknowledgements

The author declares that there is no conflict of interest that would prejudice the impartiality of this scientific work.
References


Lewis SM & Aitken RJ 2005 DNA damage to spermatozoa has impacts on fertilization and pregnancy. Cell and Tissue Research 322 33–41.

Lewis SM, Boyle PM, McKinney KA, Young JS & Thompson W 1995 Total antioxidant capacity of seminal plasma is different in fertile and infertile men. Fertility and Sterility 64 868–870.


Nallaella KP, Sharma RK, Aziz N & Agarwal A 2006 Significance of sperm characteristics in the evaluation of male infertility. Fertility and Sterility 85 629–634.


Received 4 April 2007
First decision 11 May 2007
Accepted 16 May 2007