

Perspectives

Comments on the release of the 5th edition of the WHO Laboratory Manual for the Examination and Processing of Human Semen

W. C. L. Ford

Obstetrics & Gynaecology Unit, Clinical Science at South Bristol, University of Bristol, St Michael's Hospital, Bristol BS2 8EG, UK

Abstract

The authors of the World Health Organization Semen Analysis Manual are to be congratulated on producing a new edition; it is an essential tool to disseminate good practice in andrology. However, the tests described have poor prognostic power to predict a man's fertility and show little about the underlying causes of sub-fertility. This commentary urges a revival of research into the diagnosis of male fertility. It suggests that fertility should be regarded as a continuum and that the artificial binary division between fertile and infertile should be abandoned. Models to predict a sub-fertile couple's chance of conception in a year should be developed on the basis of prospective data. These models would allow for sophisticated decision making about management. The future lies in the identification of tests to detect underlying pathologies open to specific treatment. Leads such as oxidative stress, defects in the intracellular regulation and the developing field of proteomics should be explored.

Asian Journal of Andrology (2010) 12: 59–63. doi: 10.1038/aja.2008.57

Keywords: diagnosis, male infertility, oxidative stress, predictive value of tests, sperm count

1 Introduction

The World Health Organization (WHO) manual has been a vital tool in the endeavour to achieve consistent standards of semen analysis across the world, and the authors are to be congratulated on the production of the new edition. Semen analysis is an imperfect tool but remains the cornerstone of the investigation of male infertility [1] and must be performed to a consistently high standard. The following comments are not intended to undermine the authors' achievement but to re-invigorate research efforts to improve the diagnosis of male sub-fertility.

Semen is analyzed for many reasons, but the following

comments concentrate on its role in the diagnosis of sub-fertility. In this context, the definition of normal ranges is not the best way to use information from existing tests to the highest advantage or to consider what is required of new tests. This is not to say that normal ranges have no value; they reveal whether a man's sperm count, sperm motility or sperm morphology is unusually poor and thus alert the clinical team that a 'male factor' may explain a couple's failure to conceive. Normal ranges for other components of semen, such as fructose or acid phosphatase, help to reveal problems with the accessory glands. However, they merely classify a symptom and reveal nothing about the underlying cause of the couple's failure to conceive.

2 What are we trying to achieve?

During the examination of an infertile couple, the underlying problem should be diagnosed so that it can be corrected using the appropriate treatment. In addition, a prognosis for different management regimes to guide

Correspondence to: Dr W. C. L. Ford, Obstetrics & Gynaecology Unit, Clinical Science at South Bristol, University of Bristol, St Michael's Hospital, Bristol BS2 8EG, UK.

Fax: +44-117-928-5290

E-mail: chris.ford@bristol.ac.uk

Received: 3 December 2008

Revised: 5 December 2008

Accepted: 12 December 2008



the couple to a rational choice should be provided. At present, provision of a prognosis is the dominant feature in the investigation of the male factor because very few treatments other than assisted conception (intra-uterine insemination [IUI], *in vitro* fertilization [IVF], intra-cytoplasmic sperm injection [ICSI]) have been proven to be effective. Therefore, the clinical role of semen analysis is to guide the decision of whether a couple requires treatment or should be reassured and encouraged to keep trying, and, if the former, what sort of assisted conception treatment would be appropriate. Couples should be offered the cheapest and least invasive form of treatment with an acceptable chance of success, and diagnostic procedures should provide the information on which to base this choice. We must consider the efficacy of semen analysis in this context and how it might be improved.

3 The probabilistic nature of human fertility

It is no longer sensible to treat fertility as a binary categorical variable. The very term 'infertility' is not helpful because it implies a complete incapacity to conceive, whereas most couples who fail to conceive within a year are capable of conception, just with a decreased probability of doing so in a given time [2].

The clinician must decide which of these couples is likely to conceive in an acceptable time frame. This differs from distinguishing between fertile and infertile populations because most (but not all) couples with a high chance of conception per cycle will conceive in the first year and never reach the clinic. Among couples who stopped contraception and ultimately conceived, about 30% did so in the first cycle of unprotected intercourse. After 12 months, the conception rate per cycle had declined to 10%, and the average success rate per cycle in the second year of trying was about 5% [3]. Consequently, among couples attending infertility clinics after 1 year of unprotected intercourse, even those likely to conceive will have an average spontaneous conception rate of only about 5% per cycle. Thus, examining centiles of the population with a time to pregnancy of ≤ 12 months (with an average conception rate of about 15% per cycle over that time) as in the new WHO manual or even a critical comparison of this population with non-conceivers [4] is not directly relevant to deciding which of the non-conceivers is likely to conceive in a reasonable time frame. Even prospective studies on the normal population (for example, Bonde *et al.* [5] and Zinaman *et al.* [6]) are misleading in this respect, but serve to emphasize the continuous relationship between semen quality and the chance of conception. The definition of cut-off values to define normal ranges imposes an artificial dichotomy that limits the prognostic value of the information gained from semen analysis.

These values merely classify a symptom, showing nothing about the underlying cause of the couple's failure to conceive.

In order to achieve progress, the statistical criteria to assess the efficacy of a test need to be defined and generally agreed upon. One advantage of considering fertility as a binary variable is that it permits a straightforward statistical analysis of diagnostic tests by the calculation of their sensitivity, specificity, positive and negative predictive values, overall accuracy, and likelihood ratios [7]. A few studies have applied these precepts to semen analysis (for example, Guzick *et al.* [4] and sperm function tests (for example, Polansky *et al.* [8]), but have generally found tests of male infertility to be poor discriminators. This is inevitable because the 'fertile' population will contain sub-fertile couples who have been 'lucky'; 11.4% of couples with a per-cycle conception rate of only 1% will conceive in 12 cycles, whereas the 'infertile' population will contain fertile couples who have been 'unlucky'; 7% of couples with a per-cycle conception rate of 20% will fail to conceive in 12 cycles. Therefore, any attempt to discriminate becomes a hopeless trade-off between sensitivity and specificity.

We must accept that fertility is a continuum and gear diagnostic procedures to estimate a couple's chance of conception in a reasonable period of time, such as one year [2]. These models must encompass both male and female diagnoses; a prognosis must be found for the couple and it is fruitless to consider them as independent individuals. What constitutes an acceptable chance is best decided by the couples themselves with guidance by their doctors. This decision will take into account both clinical factors, such as age and diagnosis, and the couple's feelings about delays and invasive treatments.

Treatment should be an informed choice geared to a particular couple's needs and not one based on an arbitrary division. To achieve this, we need prospective studies based on couples attending fertility clinics. Of course, such studies are not without problems. Factors such as referral bias, racial and even regional differences [9] can affect the wider application of these data; thus, verification of the model in different populations is essential. Nevertheless, some progress has been made. Investigators have been able to combine three studies from different geographical locations that all used well-established tests to produce a prognostic model for conception within 1 year [10]. This model performed well in a subsequent prospective evaluation [11]. The development and validation of these models requires advanced statistics and there must be consistency in how these are used to permit comparison between different models.

Similar issues apply to the next stage of the decision-making process—the choice among IUI, IVF and ICSI.

In Europe, ICSI is the most frequently used procedure [12], although it offers no advantage over conventional IVF in the absence of sperm dysfunction [13] and there are concerns about the health of the offspring [14]. Here again, the effect of semen quality on the outcome is a continuum. Before the introduction of ICSI, even couples with a motile normal sperm concentration < 1 million mL^{-1} had a 47% chance of obtaining four embryos, provided sufficient eggs were available (Figure 1). Since then, both ovulation induction and IVF rates have improved. An estimate of the per-cycle chance of a live birth by each technique, combined with knowledge of risks and financial cost, would allow couples to be guided to the least invasive therapy consistent with their feelings and beliefs. A crude division into normal vs abnormal fails to assist in sophisticated and comprehensive decision making.

4 What do we want from new tests?

4.1 Quality control

Whichever the method that we use to examine semen in the future, it will remain vital that the results are consistent and reliable within and between laboratories. Inconsistent results have been a long-standing problem with semen analysis. Although quality control schemes and protocols have made significant progress to address this, there still remains an unacceptable degree of variation between centers. To some extent, this is because of failure to comply with WHO protocols (for example, Riddell *et al.* [15]), but it also reflects the fact that semen

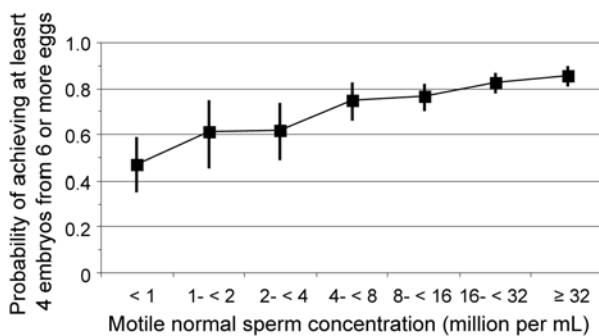


Figure 1. The probability (with 95% confidence intervals) of *in vitro* fertilization (IVF) success declines only minimally with the concentration of normal motile spermatozoa in semen. Data are from 986 consecutive first cycles of IVF in which at least six eggs were available for fertilization at the University of Bristol Centre for Reproductive Medicine before the introduction of intra-cytoplasmic sperm injection (ICSI). Obtaining four embryos is a good surrogate variable for live birth because the chance of success declines rapidly if < 4 embryos are available, but improves relatively more slowly if there are more embryos.

analysis is a skilled and partially subjective process that is inherently difficult to standardize. The answer may lie in the development of semen analysis kits that are simple to operate, for example, Fertell [16], which transfer the responsibility of quality control from many laboratories to a few factories [17]. This should be an objective for any new method, but it is important that the prognostic power of novel methods for a man's fertility be evaluated prospectively and not merely by comparison with conventional semen analysis or some other *in vitro* test.

4.2 What about sperm function tests?

It might be profitable to re-evaluate sperm function tests as indices of the chance of conception rather than as fertile or infertile discriminators. Although many sperm function tests have been able to out-perform semen analysis as predictors of both natural conception and IVF success, they have never achieved sufficient advances in discrimination to drive their incorporation as common practice. A possible exception is the post-coital test, which contributes significantly to the predictive power of the model discussed above, and which, under favorable circumstances, can greatly out-perform semen analysis [18, 19]. A major problem with the post-coital test is the difficulty of correct timing, especially in clinics that cannot be open 7 days a week. *In vitro* penetration tests, probably using mucus substitutes such as methyl cellulose or hyaluronic acid, provide an alternative that has shown promise [20–22] but require more extensive validation.

5 Diagnosis of the underlying problem

With few exceptions, we do not know why sub-fertile men become sub-fertile. Understanding this should be the primary objective of clinical andrology, because it is critical to developing both preventative strategies and informative diagnoses leading to effective therapy. Semen analysis and sperm function tests measure symptoms without revealing any underlying cause. We need diagnostic techniques that detect the underlying pathologies and lead to specific and effective treatments, rather than consider assisted conception as the only option. Although this requires significant research, there have been promising leads.

5.1 Reactive oxygen species

There is no doubt that the high production of reactive oxygen species (ROS) by crude sperm suspensions is associated with sperm dysfunction and is more prevalent in ejaculates from men who have difficulty in impregnating their partners than in the normal population. Contaminating leukocytes are the predominant source of ROS in these suspensions (see Tremellen [23]). However, although

there is a linear relationship between ROS production by seminal leukocytes and the degree of lipid peroxidation [24], whether these leukocytes infiltrate the epididymis and are responsible for the damage that occurs before ejaculation remains unknown.

Semen ROS production is an effective test to distinguish male sub-fertility patients from normal donors [25]; however, it has not been investigated in infertility patients using rigorous statistical methods. The damage done by oxidative stress depends on the balance between ROS activity and the capacity of the inter- and intra-cellular anti-oxidant defences to protect the spermatozoa. The anti-oxidant capacity of seminal plasma is deficient in some infertile men (see Tremellen [23]), making their spermatozoa more vulnerable to oxidative damage. As the impact on fertility depends on this balance and measurements of ROS are prone to artefacts that can deceive investigators (for example, Richer and Ford [26]), it may be better to base diagnostic tests on the measurement of the products of oxidative stress.

Proteins, lipids and DNA are all vulnerable to oxidative stress, but for spermatozoa, far more attention has been paid to the latter two factors. The sperm plasma membrane is rich in polyunsaturated fatty acids that are highly susceptible to peroxidation; this impairs the spermatozoon's ability to undergo acrosome reaction or penetrate zona-free hamster eggs and, at higher doses of ROS, their motility [27, 28]. Lipid peroxidation shows a highly significant log-linear relationship with the ability of spermatozoa to undergo an acrosome reaction and sperm motility [24], but has not been investigated as a diagnostic test. DNA oxidation results in strand breaks and deletions. Affected spermatozoa are less likely to fertilize than unaffected cells, but if they do, the embryo often fails to develop or may produce defective offspring. Measuring DNA oxidation is technically difficult, and the best success in terms of clinical prediction has been achieved by comparing red and green fluorescence staining of acridine orange of the sperm nucleus by flow cytometry, the sperm chromatin structure (SCSA) technique [29].

Identifying oxidative stress as an underlying cause of sperm dysfunction has the advantage that it suggests possible therapies. Administration of anti-oxidants has been attempted in several trials with mixed results; the quality of the trials has varied, but two have shown a marked improvement in fertility with treatment (see Tremellen [23]). Clearly, more work is needed. There must be a simple assay for diagnosis that can be widely used or a system for shipping samples to a limited number of expert centers. My guess is that a simple assay for lipid peroxidation could be developed, but the measurement of DNA oxidation will continue to require specialized expertise.

5.2 Defects in intracellular signalling

Another possibility is that there are defects in the signalling cascades involved in sperm maturation and capacitation, so that spermatozoa are not appropriately primed to undergo the acrosome reaction and fertilize when they reach the egg. Liu *et al.* [30, 31] reported that a common cause of fertilization failure is that spermatozoa bind to the zona pellucida but fail to undergo the acrosome reaction and penetrate it. The ARIC test may address this to some extent [32], but we must understand the defects that arise in specific regulatory cascades if we are to develop precise tests leading to specific treatments. Similar arguments apply to the regulation of sperm motility [33].

5.3 Proteomics

Finally, we must exploit the relatively new tool of proteomics. The spermatozoon's lack of active protein synthesis has meant that sperm cell biology has failed to benefit from many powerful molecular biology techniques that have led to startling advances elsewhere in biology. The intelligent application of proteomics may allow us to catch up and identify some of the molecular targets implicated in sperm dysfunction [34].

6 Conclusions

The WHO manual remains the andrologist's 'Bible' and has a vital continuing role in raising the standards of andrology laboratories. However, normal ranges are of limited value in the management of infertile couples, and semen analysis has limited prognostic power for male sub-fertility. Research to develop better diagnostic procedures has dwindled since the introduction of ICSI and must be revitalized. The key objectives are (1) to validate prospectively diagnostic models to predict the likelihood that a couple arriving at an infertility clinic will conceive naturally in the next year or that they will succeed through IUI or IVF, and (2) to build on our limited data about the causes of male infertility to devise tests that reveal the underlying pathology and suggest specific treatments.

References

- 1 Barratt CL. Semen analysis is the cornerstone of investigation for male infertility. *Practitioner* 2007; 251: 8–10, 12, 15–7.
- 2 Habbema JDF, Collins J, Leridon H, Evers JL, Lunenfeld B, *et al.* Towards less confusing terminology in reproductive medicine: a proposal. *Hum Reprod* 2004; 19: 1497–501.
- 3 Tietze C. Fertility after the discontinuation of intrauterine or oral contraception. *Int J Fertil* 1968; 13: 385–9.
- 4 Guzick DS, Overstreet JW, Factor-Litvak P, Brazil CK, Nakajima ST, *et al.* Sperm morphology, motility and concentration in fertile and infertile men. *N Engl J Med* 2001; 345: 1388–93.

- 5 Bonde JP, Ernst E, Jensen TK, Hjølland NH, Kolstad H, *et al.* Relation between semen quality and fertility: a population-based study of 430 first-pregnancy planners. *Lancet* 1998; 352: 1172–7.
- 6 Zinaman MJ, Brown CC, Selevan SG, Clegg ED. Semen quality and human fertility: a prospective study with healthy couples. *J Androl* 2000; 21: 145–53.
- 7 Greenhalgh T. How to read a paper: papers that report diagnostic or screening tests. *BMJ* 1997; 315: 540–3.
- 8 Polansky FF, Lamb EJ. Analysis of three laboratory tests used in the evaluation of male fertility: Bayes' rule applied to the postcoital test, the *in vitro* mucus migration test, and the zona-free hamster egg test. *Fertil Steril* 1989; 51: 215–28.
- 9 Auger J, Jouannet P. Evidence for regional differences of semen quality among fertile French men. *Hum Reprod* 1997; 12: 740–5.
- 10 Hunault CC, Habbema JD, Eijkemans MJ, Collins JA, Evers JL, *et al.* Two new prediction rules for spontaneous pregnancy leading to live birth among subfertile couples, based on the synthesis of 3 previous models. *Hum Reprod* 2004; 19: 2019–26.
- 11 Van der Steeg JW, Steures P, Eijkemans MJ, Habbema JD, Hompes PG, *et al.* CECERM study group (Collaborative Effort for Clinical Evaluation in Reproductive Medicine). Pregnancy is predictable: a large-scale prospective external validation of the prediction of spontaneous pregnancy in subfertile couples. *Hum Reprod* 2007; 22: 536–42.
- 12 The European IVF-monitoring programme (EIM) for the European Society of Human Reproduction and Embryology (ESHRE), Andersen AN, Gianaroli L, Felberbaum R, de Mouzon J, Nygren KG. Assisted reproductive technology in Europe, 2002. Results generated from European registers by ESHRE. *Hum Reprod* 2006; 21: 1680–97.
- 13 Bhattacharya S, Hamilton MP, Shaaban M, Khalaf Y, Seddler M, *et al.* Conventional *in-vitro* fertilisation versus intracytoplasmic sperm injection for the treatment of non-male-factor infertility: a randomised controlled trial. *Lancet* 2001; 357: 2075–9.
- 14 Zhu JL, Basso O, Obel C, Bille C, Olsen J. Infertility treatment, and congenital malformations: Danish national birth cohort infertility. *BMJ* 2006; 333: 679.
- 15 Riddell D, Pacey A, Whittington K. Lack of compliance by UK andrology laboratories with World Health Organization recommendations for sperm morphology assessment. *Hum Reprod* 2005; 20: 3441–5.
- 16 Bjorndahl L, Kirkman-Brown J, Hart G, Rattle S, Barratt CL. Development of a novel home sperm test. *Hum Reprod* 2006; 21: 145–9.
- 17 Lefevre L, Bedu-Addo K, Conner SJ, Machado-Oliveira GS, Chen Y, *et al.* Counting sperm does not add up any more: time for a new equation? *Reproduction* 2007; 133: 675–84.
- 18 Glazener CM, Kelly NJ, Weir MJ, David JS, Cornes JS, *et al.* The diagnosis of male infertility—prospective and time specific study of conception rate related to seminal analysis and post coital sperm—mucus penetration in otherwise unexplained infertility. *Hum Reprod* 1987; 2: 665–71.
- 19 Glazener CM, Ford WC, Hull MG. The prognostic power of the post-coital test for natural conception depends on duration of infertility. *Hum Reprod* 2000; 15: 1953–7.
- 20 Ivic A, Onyeaka H, Girling A, Brewis IA, Ola B, *et al.* Critical evaluation of methylcellulose as an alternative medium in sperm migration tests. *Hum Reprod* 2002; 17: 143–9.
- 21 Aitken RJ, Bowie H, Buckingham D, Harkiss D, Richardson DW, *et al.* Sperm penetration into a hyaluronic-acid polymer as a means of monitoring functional competence. *J Androl* 1992; 13: 44–54.
- 22 Neuwinger J, Cooper TG, Knuth UA, Nieschlag E. Hyaluronic acid as a medium for human sperm migration tests. *Hum Reprod* 1991; 6: 396–400.
- 23 Tremellen K. Oxidative stress and male infertility—a clinical perspective. *Hum Reprod Update* 2008; 14: 243–58.
- 24 Williams AC, Ford WC. Relationship between reactive oxygen species production and lipid peroxidation in human sperm suspensions and their association with sperm function. *Fertil Steril* 2005; 83: 929–36.
- 25 Agarwal A, Sharma RK, Nallella KP, Thomas AJ Jr, Alvarez JG, *et al.* Reactive oxygen species as an independent marker of male factor infertility. *Fertil Steril* 2006; 86: 878–85.
- 26 Richer SC, Ford WC. A critical investigation of NADPH oxidase activity in human spermatozoa. *Mol Hum Reprod* 2001; 7: 237–44.
- 27 Aitken RJ, Clarkson JS, Fishel S. Generation of reactive oxygen species, lipid peroxidation and human sperm function. *Biol Reprod* 1989; 41: 183–97.
- 28 Whittington K, Ford WC. The effect of incubation periods under 95% oxygen on the stimulated acrosome reaction and motility of human spermatozoa. *Mol Hum Reprod* 1998; 4: 1053–7.
- 29 Evenson DP, Larson KL, Jost LK. Sperm chromatin structure assay: its clinical use for detecting sperm DNA fragmentation in male infertility and comparisons with other techniques. *J Androl* 2002; 23: 25–43.
- 30 Liu DY, Baker HW. Disordered acrosome reaction of spermatozoa bound to the zona pellucida: a newly discovered sperm defect causing infertility with reduced sperm-zona pellucida penetration and reduced fertilization *in vitro*. *Hum Reprod* 1994; 9: 1694–700.
- 31 Liu DY, Clarke GN, Martic M, Garrett C, Baker HW. Frequency of disordered zona pellucida (ZP)-induced acrosome reaction in infertile men with normal semen analysis and normal spermatozoa–ZP binding. *Hum Reprod* 2001; 16: 1185–90.
- 32 Cummins JM, Pember SA, Jequier AM, Yovitch JL, Hartmann PE. A test of the human sperm acrosome following ionophore challenge (ARIC): relationship to fertility and other seminal parameters. *J Androl* 1991; 12: 98–103.
- 33 Publicover SJ, Giojalas LC, Teves ME, de Oliveira GS, Garcia AA, *et al.* Ca²⁺ signalling in the control of motility and guidance in mammalian sperm. *Front Biosci* 2008; 13: 5623–37.
- 34 Aitken RJ, Baker MA. The role of proteomics in understanding sperm cell biology. *Int J Androl* 2008; 31: 295–302.