

OPINION

Ignorance but not bliss: too little is known about the determinants of semen quality

WCL Ford

Asian Journal of Andrology (2013) 15, 174–178; doi:10.1038/aja.2012.158; published online 14 January 2013

Criticisms of the data that have been presented to demonstrate a decline in sperm counts over time reveal problems with semen analysis methods and a lack of understanding of the genetic and environmental factors that determine a man's sperm count. Potential sources of error in the WHO semen analysis protocol and some areas of ignorance about the biological and environmental factors which can influence sperm counts are briefly discussed. I conclude that there is a need to include semen analyses in a large cohort study to fill the gaps in our knowledge.

The paper by Carlsen *et al.*¹ that is being marked in this issue of *Asian Journal of Andrology* stimulated widespread concern and considerable research effort. Together with other supportive studies, e.g. Auger *et al.*,² strong evidence that testicular cancer was not only becoming more prevalent but was developing at lower ages^{3,4} and observations that other disorders of male development, e.g. cryptorchidism and hypospadias, were being more frequently reported, although this might be due to changes in clinical practice (see Thonneau *et al.*⁵), have raised serious concerns that something was going amiss with the human male. The proposition that oestrogen exposure *in utero* might be responsible by an effect on Sertoli cell development gave rise to the 'testicular dysgenesis' hypothesis to explain these effects,⁶ although the role of oestrogen has been questioned recently. As most andrologists, I have followed developments in this area with keen interest but other contributors to this issue are far more knowledgeable than me, so I intend to say nothing about the testicular dysgenesis hypothesis or the latest

consensus on whether sperm counts have indeed declined. Instead I will review some of the reasons why assessing trends in semen quality over time is difficult, and how these problems reveal gaps, or perhaps chasms, in our knowledge about the biology of human semen.

First, although there is no better alternative, the standard semen analysis is inherently imprecise and is prone to many errors. The equipment used has changed over time and with it the nature of the error, making historical comparisons difficult. The errors introduced could be random, leading to imprecision and greater difficulty in demonstrating statistical significance but without effect on the nature of the trend, or systematic, leading to bias in the results. In theory systematic error could be corrected by direct comparison of former and current techniques. However, as discussed below results can potentially be affected by technical errors whose effects are difficult to predict. These can only be detected and eliminated by good quality control systems that have only been in place for at most two decades. Therefore doubt about older data will always remain. Secondly, the environmental factors that might influence a man's sperm count have changed over time and movements of populations have changed the genetic mix in many places. Moreover, social changes might affect the characteristics of the men recruited into semen studies. How these factors affect sperm count is very poorly understood.

There are many excellent descriptions of how to do a semen analysis (e.g. Björndahl *et al.*⁷ and WHO Manual⁸), and I will not attempt to repeat these but will focus on what can so easily be done wrong in the laboratory. I will then discuss biological variation in semen quality in an attempt to highlight areas of ignorance.

METHODOLOGY

Lack of consistency

The first problem is that a number of techniques have been used and have changed over time. Changes in the type of pipette used to measure semen volumes for dilution was one of the first criticisms of the 1992 paper.⁹ Even today more than 30 years after the publication of World Health Organization (WHO) guidelines, different laboratories continue to use different techniques; a survey of laboratories participating in the German Quality Assurance Scheme found that <8% of laboratories fully followed WHO guidelines and only 55% used the recommended Neubauer chamber to determine sperm concentration.¹⁰ In a sample of 118 Chinese seminology laboratories, 49.2% relied on manual semen analysis, of which 60.2% used the recommended haemocytometer and 33.9% of laboratories used computer assisted semen analysis (CASA), with machines coming from a variety of sources.¹¹ A review concluded that there was large variation in the conduct and reporting of the test among American laboratories.¹² Differences in technique or the type of chamber used can make a big difference to the results, e.g. in our own experience of CASA, we found that the Makler chamber yielded an estimate of motile sperm concentration >30% higher than that of Microcells. Laboratories should standardize on the WHO technique, whatever their views about its merits, to achieve greater comparability of results.

Experimental errors

Even if other methods could be calibrated against the WHO procedure to account for systematic errors, there are many potential technical errors which can damage the accuracy of the analysis. Some of these pitfalls are discussed here in the hope that it will help to avoid them.

Semen is not ejaculated as a homogeneous fluid, the prostatic secretion and most of the spermatozoa being emitted before the products of the seminal vesicles. Indeed, experienced donors can split their ejaculate into five or more fractions, which have recently been more thoroughly characterized according to biochemical markers.¹³ Thus, the first problem is to ensure that the ejaculate is complete, loss of the first sperm-rich fraction being especially important. Detection of a problem has to rely on the testimony of the patient or donor. Often they may be embarrassed to report any losses, especially if not appropriately counselled about the importance of doing so and assured that no blame or shame will attach to them.

The next problem is to get a representative sample of the semen. Semen coagulates soon after ejaculation, although in general it liquefies after 10–30 min at 37 °C, occasionally liquefaction is slow or incomplete, complicating the task of taking a representative sample. Temptation to take the sample too soon because of time pressure is a potential source of error. Generally the sample removed for analysis is quite small (50 μl ⁷) thus the semen must be very thoroughly mixed to ensure that the sample is representative. This is not easy in view of the fluid's high viscosity, and ideally duplicate samples should be taken and processed rather than duplicate counts being performed on the same diluted sample. Moreover, any small error in the volume transferred can introduce a substantial percentage error. Positive-displacement pipettes have largely overcome the difficulties of pipetting viscous fluids, but air-displacement pipettes commonly used a few years ago, and various types of glass pipette used in the more distant past, are very difficult to use accurately with viscous fluids. Even with positive-displacement pipettes, it is easy to transfer several microlitres adhering to the surface of the tip unless it is carefully wiped. For diagnostic purposes where wastage is not a problem, it would be more accurate to take a larger volume for dilution since fluid adhering to the tip would be a smaller percentage of that transferred.

The semen should be diluted with a formalin-based diluent as described in the WHO guidelines. It is impossible to count moving spermatozoa accurately by eye. Again thorough mixing of the diluted semen is imperative so that the small volume that is transferred to the counting chamber will be a representative sample.

Preparing the counting chamber is a potential source of error. Even if the recommended Neubauer haemocytometer is used, regularly

calibrated with microbeads, and the coverslip pressed down until Newton's rings are seen, it is all too easy for the coverslip to rise when the sample is introduced increasing the depth of the sample and hence the sperm count. This is an even greater problem with shallower chambers as the Makler, because the percentage change in depth is larger. Moreover, although semen analysis courses run by organisations such as European Society for Human Reproduction and Embryology (ESHRE) are, it is hoped, improving the situation, laboratories do not always follow the recommended procedure for loading the chamber, or calibrate them regularly. Reproducibility might be improved if the use of disposable fixed-depth microcells were encouraged.

The 100 μm depth of the haemocytometer exceeds the focal range of most microscope objectives, so it is important that spermatozoa be allowed to settle onto the grid. This must be done in a humid chamber; if left on the bench evaporation can be a problem. If settling time is insufficient, spermatozoa in suspension will remain above the focal plane of the objective and may be overlooked. Unless spermatozoa are stained, a phase-contrast microscope is essential to visualize them easily; the thickness of the haemocytometer glass can make it difficult to focus the phase condenser properly, especially for higher power objectives. Even with the recommended settling time, some spermatozoa may stick to the coverslip and be out of the focal plane. Finally, there has to be consistency in what is considered to be a spermatozoon and counted as such. Decisions about how to treat detached heads, micro (pin) and macrocephalic spermatozoa and other unusual forms are subjective and a source of inconsistency in sperm counts. Distinguishing spermatozoa from debris is the greatest obstacle besides the expense of replacing the standard semen analysis with automated methods or in establishing them as a 'gold standard'.

The absence of a consistent approach on how the above pitfalls can be avoided makes the validity of comparisons of semen analysis results between centres and over time difficult to establish beyond doubt, since the nature of the errors is unknown. This is disappointing since clear guidelines have been available for a long time. Moreover, most semen analyses are performed as fertility tests; under pressure in a busy laboratory, there may be a temptation to believe that great accuracy is not vitally important, particularly when the sperm count is clearly more than adequate. Duplicate measurements would be a more effective check for error if the whole process,

from taking and diluting the sample to counting, was duplicated rather than just making two counts on the same dilution. Errors in the estimation of semen volume can be considerable¹⁴ and will contribute to errors in calculation of the total sperm count. A significant improvement in the consistency of measurements of sperm concentration can be achieved by central processing of semen samples preserved in formalin.¹⁵ This method has been employed in a clinical trial of a male contraceptive.¹⁶ This system could be used to ensure consistent results in future epidemiological studies.

Interejaculate variation

Extremely wide variation in the quality of different ejaculates from the same man is well known and is illustrated in the WHO guidelines for semen analysis.⁸ The only solution is to base any population statistics on sufficiently large numbers. However, large numbers of subjects are required to gain sufficient statistical power to detect small changes and it may be difficult to recruit sufficient men. How semen quality is affected by the frequency of ejaculation is also pertinent since sexual habits have changed over time.¹⁷ Cooper *et al.*¹⁸ have shown that up to about 7 days sperm count increases with abstinence time (although motility may decline after a shorter period) thus increased sexual activity could account for a decline in sperm concentration in recent times.

Statistics

The first statistical consideration in considering measurement of sperm concentration is that sufficient spermatozoa should be counted. The count follows a Poisson distribution, so that the standard deviation is equal to the square root of the mean. Thus, if 100 sperm cells are counted, the standard deviation is 10 (10%) or for 400 sperm cells 20 (5%) and these need to be multiplied by 1.96 to calculate the 95% confidence limits. Thus sufficient spermatozoa must be counted to achieve sufficient precision. It would be good to give 95% confidence intervals when reporting sperm counts to expose this inevitable uncertainty. Greater precision can only come from counting more spermatozoa. It is too time-consuming to count >400 spermatozoa manually in a service laboratory and automated procedures need to be developed. Cytometric methods based on DNA staining to identify haploid cells, together with light scattering measurements for shape discrimination, should be capable of good accuracy¹⁹ and have been shown to be effective.¹⁵

In summarizing the results, it must be borne in mind that the distribution of sperm counts in the population is highly skewed. Thus, the mean and standard deviation give a misleading picture of the central tendency and dispersion of the population. The data should be presented using non-parametric measures (median, interquartile range) or transformed for analysis; a cube root transformation is generally held to give the best fit to a normal distribution.²⁰

When comparing population values, the most important consideration is that the samples are taken at random to eliminate bias. This raises difficulties when only a small fraction of those approached agree to take part, since it is likely that those who volunteer will be different from those who decline. For example, volunteers may be better-educated or have some reason, such as concern about fertility, to be interested in the results. Bias can also be introduced by recruitment from a restricted group within the population, e.g. those attending hospital. Doubt about whether the samples of the population used in historic samples are truly random and representative is another major problem in deciding if trends over time are real. Furthermore, in some studies, bias was introduced by a misguided attempt to focus on the normal population by omitting counts below a certain value,²¹ which is exacerbated if the definition of normality changes over time.

When comparing populations over time, we have to consider factors that may have altered the nature of the sample population over time, and so introduce bias, as well as changes in the environment that might have altered sperm counts.

Biological factors that could cause bias

The attempt to establish if the decline in sperm counts over time that some studies have reported is a true deterioration or an artefact due to other changes in the population under scrutiny has revealed how little we know about the long- and short-term factors that determine a man's sperm count. Some of the major factors involved are discussed below.

Age. Since more couples than in the past are choosing to postpone childbearing until later in life, and fertility problems are the most common reasons to have a semen analysis, it is possible that men examined recently are on average older than those studied some time ago. The effect of age on male fertility and semen quality remains controversial. There are well-known examples of men, e.g.

Charlie Chaplin, fathering children while in their 70s; however, our epidemiological study suggested that male fertility declines slowly as men become >35 years old.²² Although the main model in this paper has been correctly criticized for being based on age at the time of conception, and not at the start of the attempt to conceive,²³ the alternative model based on the difference between the age of the women and her husband remains valid. Similar conclusions were reached by Dunson *et al.*,²⁴ on the basis of 782 couples using natural family planning methods. Some of the decrease in fertility can be attributed to reduced *libido* and hence coital frequency, perhaps associated with erectile dysfunction (see Sartorius *et al.*²⁵ and Harris *et al.*²⁶). There is conflicting evidence about the effect of advancing age on semen quality. Some authors report that sperm concentration or the total per ejaculate declines.^{27–32} However, other studies report that sperm concentration remains stable while motility and morphology decline.³³ In a review of the literature, Kidd *et al.*³⁴ also concluded that semen volume, sperm motility and morphology but not sperm concentration decreased with age. Obesity, which tends to be more prevalent in older men is a potential confounding factor although a meta-analysis found no association between body mass index and semen quality.³⁵ Apart from the impact of age on the 'decline' question, it is an important issue in view of peoples' desire to start their families when older. It also resonates with the idea that free radical damage to sperm DNA increases with advancing age, leading to a related increase in the risk of genetic damage, as well as to decreased fertility (see Sartorius *et al.*²⁵ and Desai *et al.*³⁶). Here again there are huge gaps in our knowledge; the source of free radicals in the testis and epididymis is unknown, as are the defence mechanisms and why they can be overwhelmed: this will be difficult to sort out, as even the source of reactive oxygen species in sperm suspensions remains unresolved.

An area of still greater ignorance is how a man's semen quality changes during his lifetime. Has the oligozoospermic man who presents at the infertility clinic in his mid-thirties had a low sperm count since puberty, or has it declined gradually over the years or abruptly, as a result of some unrecognized disease or trauma, or perhaps do all three scenarios occur in different subjects? Do the men who present with poor semen quality and difficulty in initiating their first pregnancy, but who are not seen again because subsequent children come without difficulty, experience an improvement in semen quality or just get

lucky? Other than short-term studies suggesting that severely oligozoospermic men are at risk of developing azoospermia,³⁷ and comparison of a man's initial sample with a later one closer to the conception of a spontaneous pregnancy,³⁸ I know of no studies of semen quality on large numbers of men at regular intervals throughout their reproductive life-span. I suspect that semen quality will follow a wide range of time-courses across the population but until we know what these are, we lack vital information to deduce the multiple causes of poor semen quality.

Ethnic and regional factors. The paper by Auger *et al.*² was one of the more convincing studies suggesting that sperm counts have decreased, since it was based on results from semen donors analysed in a single laboratory in Paris using consistent semen analysis and recruitment techniques. However, this was criticized because of possible changes in the nature of the donor population due to migration into and out of the city.

A number of studies suggest that there is considerable regional variation in semen quality, exceeding the change attributable to declining sperm counts. There is variation in semen between different regions of France,³⁹ and different European cities (Copenhagen, Paris, Edinburgh and Turku),⁴⁰ although these differences were not correlated with time to pregnancy in these cities,⁴¹ the Baltic states⁴² and the United States.⁴³ It is not yet clear whether these differences are due to genetic or to environmental differences and recruitment bias is difficult to exclude.

A more contentious issue is whether sperm count varies between races, as this knowledge is important to set appropriate reference ranges for clinical purposes and to allow population data to be interpreted against a background of changing ethnic mixes. It might also shed some light on the effect of different cultural habits, e.g. diet, extended breast-feeding and on genetic factors. There is rather little good information about this topic because of differences in the type of population under study, the analytical technique employed and the statistical treatment of results. These limitations are shared by the studies cited below. A study of fertile men from five Japanese cities, done with the same protocol as the European studied mentioned above, concluded that sperm concentrations in Japanese men were similar to those in the worst European centre (Copenhagen), although the inverse correlation between sperm count and the incidence of testicular

cancer seen in Europe was not observed in Japan.⁴⁴ Another study in Japan suggested that the sperm concentration might be higher, but took no account of the severe skewing of sperm count data in its statistical analysis.⁴⁵ A recent study in south-west China concluded that sperm counts among the general population of men there were higher than those in the United States or the Baltic states but lower than those in France; however, concentrations in this study were higher than those in other reports from China.⁴⁶

ENVIRONMENTAL FACTORS THAT MIGHT AFFECT SPERM COUNT

Occupation, lifestyle and chemical exposures

This is a vast topic too broad to review here, and others have greater expertise to do it, but in the modern age from the cradle to the grave, we are exposed to a broad cocktail of chemicals, the nature of the cocktail depending, among other things, on whether we live in an urban or rural environment, our lifestyle choices and the regulatory policies of the society in which we live.

Disruption of testicular endocrinology *in utero* was a key factor in the testicular dysgenesis theory propounded by Sharpe and Skakkebaek.⁶ There is now some evidence that endocrine disrupters can accumulate in the mother's body and affect reproductive functions; their presence in follicular fluid is associated with decreased competence of oocytes,⁴⁷ occupational exposures to chemicals affect foetal growth,⁴⁸ and high levels of organo-halogen endocrine disrupters in maternal serum at 35 weeks of pregnancy correlate with changes in sex hormone concentrations and testicular size in boys up to 18 months of age.⁴⁹ These reports illustrate the growing amount of *in vitro* and animal evidence that such effects are possible (see Phillips *et al.*⁵⁰) and have had an impact on human male fertility (see Sikka *et al.*⁵¹).

There are numerous historical examples of adult chemical exposures that have affected sperm counts and male fertility, e.g. the drug sulphasalazine used to treat ulcerative colitis affected semen quality and fertility;⁵² occupational exposure to dibromo-chloropropane (DBCP) made male workers infertile (see Potashnik *et al.*⁵³) and exposure to gossypol as a contaminant of cooking oil in China caused male infertility, and led to its investigation as a male contraceptive lead (see Prasad *et al.*⁵⁴). There has been concern about solvents, notably glycol ethers (e.g. Cherry *et al.*⁵⁵) and pesticides (see Hanke *et al.*⁵⁶). Lifestyle factors such as smoking, alcohol

and caffeine intake and stress have also been associated with lowering of semen quality^{57–60}, although better evidence is needed to confirm these effects.⁶¹

Determining exposure is another difficult area because good quantitative exposure data are difficult to collect on large numbers of men (exposures *in utero* are even harder to measure) and epidemiological studies often rely on occupation or geographical location known to be associated with given exposures. Most of all it is difficult because men are exposed to very many chemicals which interact in unknown ways—what is the effect of being exposed to several hundred chemicals each at a small fraction of their toxic dose—do they act independently, additively or synergistically? The contribution of the toxicity of manmade chemicals to any decline in sperm count remains an open question.

Temperature

The sensitivity of spermatogenesis to heat is long established (see Setchell⁶²) and insulating underwear to raise testicular temperature has even been proposed as a means of contraception.⁶³ There is also well-recognized seasonal variation in sperm counts which are significantly higher in the winter, as demonstrated in the European study cited above.⁴⁰ Since 1940, the average worldwide surface temperature has risen by almost 0.6°C⁶⁴ and more people live in cities: an urban area with a population of 1 million is estimated to be 1–3 °C hotter than the surrounding countryside.⁶⁵ In Europe and North America, homes and workplaces are better heated than in the past and there has been a tendency towards the wearing of tighter underpants (Y fronts). Taking these trends together, it is possible that increased testicular temperature could have contributed to a decline in sperm counts.

Conclusions—what needs to be done

There is nothing we can do about how semen was assessed in the past, and retrospective comparisons will always be difficult; however, we must all throw our weight behind getting every laboratory to standardize on the WHO protocol. The protocol itself should be scrutinized to eliminate every source of error possible, e.g. the introduction of factory-prepared fixed depth slides would decrease the risk of errors due to poor placing of the haemocytometer coverslip. Every laboratory should calibrate its equipment regularly. The adoption of cytometric methods to provide a 'gold standard' for sperm counting would be valuable and should be possible using DNA stains to distinguish haploid cells

and size gating, possibly combined with cell surface markers to eliminate inclusion of immature germ cells. The ability to count 10 000 cells would decrease the standard deviation to 1%. The feasibility of this approach is illustrated by the study of Jonckheere *et al.*¹⁵

While case control studies, e.g. Chemicals and pregnancy study (CHAPS)⁵⁵ and cross-sectional studies (e.g. Hansen *et al.*⁵⁹) are valuable, the best way to examine the influence of environmental and genetic factors on sperm count and semen quality, while getting information on the development of these factors over a man's lifetime, is to incorporate semen analyses into a large cohort study. Some of these gather extremely detailed data on the gestation and upbringing of the child as well as medical and exposure information and material for genetic evaluation from the children, parents and grand-parents. Examples from the United Kingdom include the Avon Longitudinal Study of Parents and Children,^{66,67} which is continuing since its initiation in 1991 and has enrolled more than 14 000 mothers; the Millennium Cohort Study, which is following 19 000 children born in 2000–2001,⁶⁸ and the UK Birth Cohort Study, which aims to collect data on 100 000 children across the United Kingdom.⁶⁹ The unbiased recruitment of healthy men to provide semen samples for research is difficult whatever the study design, but participants in these large cohort studies are often very committed to the study and have a trusting relationship with the staff. These factors should make it possible to achieve a satisfactory level of recruitment to an 'add-on' semen study. The Avon Longitudinal Study of Parents and Children cohort in particular is an opportunity waiting to be seized and could rapidly produce data about the factors associated with a young man's semen quality although looking at the evolution of sperm counts over time is clearly a long-term endeavour.

COMPETING FINANCIAL INTERESTS

The author declares that there are no competing financial interests.

- 1 Carlsen E, Giwercman A, Keiding N, Skakkebaek NE. Evidence for decreasing quality of semen during past 50 years. *BMJ* 1992; **305**: 609–13.
- 2 Auger J, Kunstmann JM, Czyglik F, Jouannet P. Decline in semen quality among fertile men in Paris during the past 20 years. *N Engl J Med* 1995; **332**: 281–5.
- 3 Osterlind A. Diverging trends in incidence and mortality of testicular cancer in Denmark, 1943–1982. *Br J Cancer* 1986; **53**: 501–5.
- 4 Hugue E, Matsuda T, Thonneau P. Increasing incidence of testicular cancer worldwide: a review. *J Urol* 2003; **170**: 5–11.

- 5 Thonneau PF, Candia P, Mieuisset R. Cryptorchidism: incidence, risk factors, and potential role of environment; an update. *J Androl* 2003; **24**: 155–62.
- 6 Sharpe RM, Skakkebaek NE. Are oestrogens involved in falling sperm counts and disorders of the male reproductive tract? *Lancet* 1993; **341**: 1392–5.
- 7 Björndahl L, Motimer D, Barratt CLR, Castilla JA, Menkveld R *et al*. A Practical Guide to Basic Laboratory Andrology. Cambridge: Cambridge University Press; 2010.
- 8 WHO Department of Reproductive Health and Research. WHO Laboratory Manual for the Examination and Processing of Human Semen. 5th edition. Geneva: World Health Organization; 2010.
- 9 Tummon IS, Mortimer D. Decreasing quality of semen. *Br Med J* 1992; **305**: 1228–9.
- 10 Mallidis C, Cooper TG, Hellenkemper B, Lablans M, Uckert F *et al*. Ten years' experience with an external quality control program for semen analysis. *Fertil Steril* 2012; **98**: 611–6.
- 11 Lu JC, Zhang HY, Hu YA, Huang YF, Lü NQ. A survey on the status of semen analysis in 118 laboratories in China. *Asian J Androl* 2010; **12**: 104–10.
- 12 Keel BA. How reliable are results from the semen analysis? *Fertil Steril* 2004; **82**: 41–4.
- 13 Ndovi TT, Parsons T, Choi L, Caffo B, Rohde C *et al*. A new method to estimate quantitatively seminal vesicle and prostate gland contributions to ejaculate. *Br J Clin Pharmacol* 2007; **633**: 404–20.
- 14 Cooper TG, Brazil C, Swan SH, Overstreet JW. Ejaculate volume is seriously underestimated when semen is pipetted or decanted into cylinders from the collection vessel. *J Androl* 2007; **28**: 1–4.
- 15 Jonckheere J, Callewaert N, Grootenhuis AJ, Kersemaekers WM, Cooper TG. Feasibility study for the centralized measurement of sperm concentration. *Int J Androl* 2005; **28**: 311–5.
- 16 Mommers E, Kersemaekers WM, Ellicesen J, Kepers M, Apter D *et al*. Male hormonal contraception: a double-blind, placebo-controlled study. *J Clin Endocrinol Metab* 2008; **93**: 2572–80.
- 17 Menger HJ. Sexual revolution and sperm count. *Brit Med J* 1994; **308**: 1440–1.
- 18 Cooper TG, Keck C, Oberdieck U, Nieschlag E. Effects of multiple ejaculations after extended periods of sexual abstinence on total, motile and normal sperm numbers, as well as accessory gland secretions, from healthy normal and oligozoospermic men. *Hum Reprod* 1993; **8**: 1251–8.
- 19 Ferrara F, Daverio R, Mazzini G, Bonini P, Banfi G. Automation of human sperm cell analysis by flow cytometry. *Clin Chem* 1997; **43**: 801–7.
- 20 Handelsman DJ. Optimal power transformations for the analysis of sperm concentration and other semen variables. *J Androl* 2002; **23**: 629–34.
- 21 Bromwich P, Cohen J, Stewart I, Walker A. Decline in sperm counts: an artefact of changed reference range of "normal"? *Br Med J* 1994; **309**: 19–22.
- 22 Ford WC, North K, Taylor H, Farrow A, Hull MG *et al*. The ALSPAC Study Team (Avon Longitudinal Study of Pregnancy and Childhood). Increasing paternal age is associated with delayed conception in a large population of fertile couples: evidence for declining fecundity in older men. *Hum Reprod* 2000; **15**: 1703–8.
- 23 Sallmén M, Luukkonen R. Is the observed association between increasing paternal age and delayed conception an artefact? *Hum Reprod* 2001; **16**: 2027–8.
- 24 Dunson DB, Colombo B, Baird DD. Changes with age in the level and duration of fertility in the menstrual cycle. *Hum Reprod* 2002; **17**: 1399–403.
- 25 Sartorius GA, Nieschlag E. Paternal age and reproduction. *Hum Reprod Update* 2010; **16**: 65–79.
- 26 Harris ID, Fronczak C, Roth L, Meacham RB. Fertility and the aging male. *Rev Urol* 2011; **13**: e184–90.
- 27 Ng KK, Donat R, Chan L, Lalak A, di Pierro I *et al*. Sperm output of older men. *Hum Reprod* 2004; **19**: 1811–5.
- 28 Hellstrom WJ, Overstreet JW, Sikka SC, Denne J, Ahuja S *et al*. Semen and sperm reference ranges for men 45 years of age and older. *J Androl* 2006; **27**: 421–8.
- 29 Levitas E, Lunenfeld E, Weisz N, Friger M, Potashnik G. Relationship between age and semen parameters in men with normal sperm concentration: analysis of 6022 semen samples. *Andrologia* 2007; **39**: 45–50.
- 30 Cardona MW, Berdugo J, Cadavid Jaramillo A. The effects of male age on semen parameters: analysis of 1364 men attending an andrology center. *Aging Male* 2009; **12**: 100–3.
- 31 Paasch U, Grunewald S, Kratzsch J, Glander HJ. Obesity and age affect male fertility potential. *Fertil Steril* 2010; **94**: 2898–901.
- 32 Molina RI, Martini AC, Tissera A, Olmedo J, Senestrari D *et al*. [Semen quality and aging: analysis of 9,168 samples in Cordoba, Argentina.] *Arch Esp Urol* 2010; **63**: 214–22.
- 33 Mukhopadhyay D, Varghese AC, Pal M, Banerjee SK, Bhattacharyya AK *et al*. Semen quality and age-specific changes: a study between two decades on 3,729 male partners of couples with normal sperm count and attending an andrology laboratory for infertility-related problems in an Indian city. *Fertil Steril* 2010; **93**: 2247–54.
- 34 Kidd SA, Eskenazi B, Wyrobek AJ. Effects of male age on semen quality and fertility: a review of the literature. *Fertil Steril* 2001; **75**: 237–48.
- 35 MacDonald AA, Heribson GP, Showell M, Farquhar CM. The impact of body mass index on semen parameters and reproductive hormones in human males: a systematic review with meta-analysis. *Hum Reprod Update* 2010; **16**: 293–311.
- 36 Desai N, Sabanegh E Jr, Kim T, Agarwal A. Free radical theory of aging: implications in male infertility. *Urology* 2010; **75**: 14–9.
- 37 Bak CW, Song SH, Yoon TK, Lim JJ, Shin TE *et al*. Natural course of idiopathic oligozoospermia: comparison of mild, moderate and severe forms. *Int J Urol* 2010; **17**: 937–43.
- 38 Juárez Bengoa A, Fernández Larios JP, Rojas Ruiz JC, Silvestre Tomassoni JR, Villanueva Diaz CA *et al*. [Changes in a semen analysis related to spontaneous pregnancy in patients with infertility history]. *Ginecol Obstet Mex* 2006; **74**: 48–54.
- 39 Auger J, Jouannet P. Evidence for regional differences of semen quality among fertile French men. *Hum Reprod* 1997; **12**: 740–5.
- 40 Jørgensen N, Andersen AG, Eustache F, Irvine DS, Suominen J *et al*. Regional differences in semen quality in Europe. *Hum Reprod* 2001; **16**: 1012–9.
- 41 Jensen TK, Slama R, Ducot B, Suominen J, Cawood EH *et al*. Regional differences in waiting time to pregnancy among fertile couples from four European cities. *Hum Reprod* 2001; **16**: 2697–704.
- 42 Punab M, Zilaitiene B, Jørgensen N, Horte A, Matulevicius V *et al*. Regional differences in semen qualities in the Baltic region. *Int J Androl* 2002; **25**: 243–52.
- 43 Swan SH, Brazil C, Drobnis EZ, Liu F, Kruse RL *et al*. Geographic Differences in Semen Quality of Fertile U.S. Males. *Environ Health Perspect* 2003; **111**: 414–20.
- 44 Iwamoto T, Nozawa S, Yoshiike M, Hoshino T, Baba K *et al*. Semen quality of 324 fertile Japanese men. *Hum Reprod* 2006; **21**: 760–5.
- 45 Baba K, Nishida T, Yoshiike M, Nozawa S, Hoshino T *et al*. Current status of reproductive function in Japanese fertile men: international collaborative project on a study of partners of pregnant women. *Int J Androl* 2000; **23**(Suppl 2): 54–6.
- 46 Li Y, Lin H, Ma M, Li L, Cai M *et al*. Semen quality of 1346 healthy men, results from the Chongqing area of southwest China. *Hum Reprod* 2009; **24**: 459–69.
- 47 Petro EM, Leroy JL, Covaci A, Franssen E, de Neubourg D *et al*. Endocrine-disrupting chemicals in human follicular fluid impair *in vitro* oocyte developmental competence. *Hum Reprod* 2012; **27**: 1025–33.
- 48 Snijder CA, Roelveland N, te Velde E, Steegers EA, Raat H *et al*. Occupational exposure to chemicals and fetal growth: the Generation R Study. *Hum Reprod* 2012; **27**: 910–20.
- 49 Meijer L, Martijn A, Melessen J, Brouwer A, Weiss J *et al*. Influence of prenatal organohalogen levels on infant male sexual development: sex hormone levels, testes volume and penile length. *Hum Reprod* 2012; **27**: 867–72.
- 50 Phillips KP, Tanphaichitr N. Human exposure to endocrine disruptors and semen quality. *J Toxicol Environ Health* 2008; **11**: 188–220.
- 51 Sikka SC, Wang R. Endocrine disruptors and estrogenic effects on male reproductive axis. *Asian J Androl* 2008; **10**: 134–45.
- 52 Levy AJ, Fisher AM, Hughes L, Hendry WF. Sulphasalazine and male infertility. *Lancet* 1979; **314**: 276.
- 53 Potashnik G, Porath A. Dibromochloropropane (DBCP) a 17 year reassessment of testicular function and reproductive performance. *J Occup Environ Med* 1995; **37**: 1287–92.
- 54 Prasad MR, Diczfalusy E. Gossypol. *Int J Androl* 1982; (Suppl 5): 53–70.
- 55 Cherry N, Moore H, McNamee R, Pacey A, Burgess G *et al*. Occupation and male infertility: glycol ethers and other exposures. *Occup Environ Med* 2008; **65**: 708–14.
- 56 Hanke W, Jurewicz J. The risk of adverse reproductive and developmental disorders due to occupational pesticide exposure: an overview of current epidemiological evidence. *Int J Occup Environ Health* 2004; **17**: 223–43.
- 57 Ramlau-Hansen CH, Thulstrup AM, Aggerholm AS, Jensen MS, Toft G *et al*. Is smoking a risk factor for decreased semen quality? A cross-sectional analysis. *Hum Reprod* 2007; **22**: 188–96.
- 58 Jensen TK, Swan SH, Skakkebaek NE, Rasmussen S, Jørgensen N. Caffeine intake and semen quality in a population of 2,554 young Danish men. *Am J Epidemiol* 2010; **171**: 883–91.
- 59 Hansen ML, Thulstrup AM, Bonde JP, Olsen J, Håkonsen LB *et al*. Does last week's alcohol intake affect semen quality or reproductive hormones? A cross-sectional study among healthy young Danish men. *Reprod Toxicol* 2012; **34**: 457–62.
- 60 Li Y, Lin H, Li Y, Cao J. Association between socio-psycho-behavioral factors and male semen quality: systematic review and meta-analyses. *Fertil Steril* 2011; **95**: 116–23.
- 61 Sardeu JC, Hughes CL, Agarwal S, Foster WG. Alcohol, drugs, caffeine, tobacco, and environmental contaminant exposure: reproductive health consequences and clinical implications. *Crit Rev Toxicol* 2010; **40**: 633–52.
- 62 Setchell BP. The Parkes Lecture. Heat and the testis. *J Reprod Fertil* 1998; **114**: 179–94.
- 63 Mieuisset R, Bujan L. The potential of mild testicular heating as a safe, effective and reversible contraceptive method for men. *Int J Androl* 1994; **17**: 186–91.
- 64 NASA. 2012. <http://climate.nasa.gov/keyIndicators/#globalTemp> (visited 29/06/2012).
- 65 EPA. 2012. <http://www.epa.gov/heatIslD/about/index.htm> (visited 29/06/2012).
- 66 Golding J. 'Children of the nineties. A longitudinal study of pregnancy and childhood based on the population of Avon (ALSPAC). *West Engl Med J* 1990; **105**: 80–2.
- 67 University of Bristol. 2012. <http://www.bristol.ac.uk/alspac/> (visited 29/06/2012).
- 68 Centre for Longitudinal Studies. 2012. <http://www.cls.ioe.ac.uk/page.aspx?&siteid=851&siteid=Welcome+to+the+Millennium+Cohort+Study> (visited 29/06/2012).
- 69 University College London. 2012. <http://www.ucl.ac.uk/ich/research-ich/mrc-cech/cohort-studies/uk-bcs> (visited 29/06/2012)