

Commentary

Afterword to *Semen Analysis in 21st Century Medicine* special issue in *Asian Journal of Andrology*

David J. Handelsman^{1*}, Trevor G. Cooper^{2*}

¹*Reproductive Endocrinology & Andrology Director, ANZAC Research Institute, Sydney, Australia*

²*Andrology Laboratory, Institute of Reproductive Medicine of the University, Munster, Germany*

Asian Journal of Andrology (2010) 12: 118–123. doi: 10.1038/aja.2009.50

We summarise and comment on the topics discussed by the contributors to this Special Issue.

As an introduction, Ann Jequier's [1] generally positive assessment of the new-look manual was tempered with her reservations about reference values, on which other chapters dwelt at length. Aspects of laboratory technique were discussed by Charlene Brazil [2], whose offering gives a rare insight into the actual workings of andrology laboratories and highlights the need for greater involvement of laboratory directors (often clinicians) in assessing the quality of data provided. Experience with the German external quality control programme for semen analysis confirms that some laboratories treat the quality control (QC) samples differently (using a Neubauer chamber, for instance) from routine samples (a Makler chamber) so that the exercise may be undermined and fail to provide proper control of the lab's quality of work. It is one thing to have QC accepted in principle but the essential job of "policing" semen analysis QC falls in no-one's court. Allan Pacey's [3] contribution supports wholeheartedly quality assurance (QA) in andrology but also zeros in on a potential quantitative problem in the precision of normal sperm morphology assessment, when such values are low.

Rune Eliasson's [4] characteristically blunt contribution lambasts the whole process of scientific consensus, the mainstay of world health organization (WHO) committees' recommendations. He considers that the perspective of the handbook are too narrow; its recommendations on assessing sperm morphology are inadequate; there to be insufficient emphasis on interpretation of results (cumulative importance of laboratory findings, interpretation of reference limits) or on manipulation of data (e.g. dividing total sperm numbers by days of abstinence and the testicular volume) and finds no mention of how to measure testicular volume. He laces this with invective against the lack of scientific method of those working in ART and reserves particular venom for the morphology method espoused in WHO3 [5] and WHO4 [6] that is given even more prominence in WHO5 [7].

Once our hackles are down again, though, his view contain some prickly wisdom. First, if only normal forms are examined, the analysis could provide little indication of testicular dysfunction (other than the reduced output of normal forms) that may be indicated by production of particular abnormal forms. Second, as the normative reference values are so low, it can provide little scope for diagnosis, making it a meaningless end-point.

Eliasson's [4] suggestion is that different criteria for categorising spermatozoa as morphologically normal (borderline forms considered normal, as in earlier editions of the WHO manual) should be implemented. These should provide higher percentages of normal forms and thus increase the possibility of distinguishing men with lower percentages of such forms. Thus far more workable and informative reference limits around 40%–60% would be provided by using less restrictive assessment. However, unless good evidence supports the criteria that determine whichever morphological criteria are chosen as reference threshold, it seems to be putting the cart before the horse; choosing an end-point for its ability to distinguish subgroups rather than to separate groups on biological function that comprise authentic clinical fertility, rather than laboratory surrogate, endpoints. The restriction to artificial IVF, rather than natural fertility populations, of the supportive science for the more restrictive definitions is also germane to Eliasson's criticism [4].

His criticism of the human inability to decide all-or-nothing (normal or abnormal) morphology is illustrated by a "brain

*Guest-Editors:

Prof. David J. Handelsman Email: djh@anzac.edu.au

Dr Trevor G. Cooper Email: tgcooper@gmx.de



teaser” optical illusion, which, however, is even more applicable to the mind games involved in defining sperm heads as cigar forms, pear- or lemon-shaped, which would be required to control the assessment within and among laboratories of all the abnormal morphological forms that he suggests be analysed.

Both Eliasson [4] and Lars Björndahl [8] consider that abandoning the distinction between fast- and slow-progressing spermatozoa is a step backwards. Björndahl’s [8] contribution refutes WHO5’s line of thinking as unnecessarily defeatist, although Pacey [3] confirms the WHO viewpoint that agreement between technicians in velocity assessment is poor. Björndahl [8] makes the valid points that (1) as motility is essential for so many aspects of sperm function, it is logical to distinguish between fast and slow spermatozoa; (2) doing so is no more subjective than is assessing morphology, which is stressed in WHO5, and (3) QC can be used to improve technician evaluation of motility, just as it can for morphology. Furthermore, he cites two papers (but only two!) that provide evidence that assessing “grade a” spermatozoa, either alone or together with results of zona binding, can be used to predict IVF success and supplies references to support the view that the velocity of 25 $\mu\text{m/s}$ for a “grade a” spermatozoon was based on kinematic data. Again the restriction to highly artificial IVF setting rather than the healthy population is a notable limitation.

Examination of the references (some of which are reviews and cite previous publications), however, reveals that most do not mention this value at all and, when mentioned, that it was obtained from different sorts of men, by the use of different chambers and different sperm selection techniques than those techniques used to provide the other reference data provided in WHO5 [7]. The decision to reject categorising progressive spermatozoa into fast and slow in WHO5 was also based on the inability of technicians to gauge velocities accurately and that if velocities are necessary, CASA systems should therefore provide them. Interestingly, when the speed of a technicians’ grade “a” spermatozoa was first estimated, it was about 60 $\mu\text{m/s}$ [9]. If technicians can indeed reproducibly tell spermatozoa swimming at 60 $\mu\text{m/s}$ apart from far slower ones (25 $\mu\text{m/s}$?) then Björndahl’s idea may well be worth taking up; backed up, as he recommends, by internal and external QC.

Despite its not being spelled out in WHO5, there is no reason why laboratories should not attempt to monitor sperm velocities (objectively when possible). Indeed, perhaps data on the percentages of fast and slow spermatozoa should be collected, but this will only be achieved with consistent agreement on velocities among technicians. When sufficient data are accumulated for critical evaluation, it could make an evidence-based return in WHO6. Scatter analysis of kinematic data from CASA systems may be even better able to distinguish subgroups of men from their semen analysis.

Eliasson’s comments [4] on sperm morphology are also taken up, but more politely, by Jaques Auger [10]. First, he returned to the fact that there is little room for any group with a lower percentage of normal forms to be distinct from the fertile men, which naturally reduces the diagnostic value of the reference limit. If men can be fertile with such a low percentage of such normal cells, then indeed perhaps it is not a parameter worth examining. However, whether the actual percentage *per se* is useful is questionable: the likely more relevant parameter is the total number of morphologically normal spermatozoa in the ejaculate and men with even lower percentages of normal forms than fathers may well have far fewer total spermatozoa as well. Overlap of these derived values from fertile and infertile men may also occur but James Boyd [11] presented one way of analysing distributions of data that overlap significantly.

Augers’s [10] thoughtful contribution provides an overview of a more comprehensive sperm morphology systems he has developed to advantage. He mentions prognosis from low lower reference limits; argues that defining abnormal sperm forms may be useful for diagnosis, prognosis and research purposes; that assessment of multiple defect indices may be useful when assessing various environmental factors or pathological conditions and that computer-aided sperm morphologists (CASMA) systems are optimal for reproductive toxicology and basic research. All these categorizations and claims, however, need formal prospective testing for their utility, ideally in natural fertility populations rather than merely in IVF or laboratory settings.

The essential differences between the eye-brain system of the technician (good for pattern recognition but not measurement) and CASMA (good for measurement but not pattern recognition) are also discussed. Whether the same problems besetting CASA (the inability of technicians to reproduce computer results, e.g. determine velocity and hence agree on categorization of grade “a” spermatozoa) will beset CASMA (the inability of the CASMA to mimic the technician and focus through a sample to distinguish cell from background) seems highly likely. Perhaps their use will be limited to the use of external quality control programmes in setting target values objectively.

Roelof Menkveld [12] presents the WHO5-supported concept that has initiated such heated discussion from Eliasson [4] and Auger [10]. Acceptance of the strict categorisation of normal morphological forms (borderline forms considered abnormal) lies in the fact that they represent a selected, potentially fertilising, population of spermatozoa that have successfully penetrated cervical mucus and that display no major defects of the head or tail. A similar uniformity in sperm morphology is seen in the cells selectively bound to the zona pellucida during attempts at IVF. Populations of such spermatozoa are found in semen but represent a minority of the cell types. Menkveld [12] argues that these values agree, not unexpectedly, with

other reports where the strict categorisation of normal forms has been applied, but he cautions against a tendency to become so strict that no normal forms are ever seen. As Brazil [2] pointed out, distinguishing between normal and borderline and abnormal is not easy and this may be where CASMA may play an important role. Downloading images to a centralised computer for analysis is not a far-fetched idea.

Although such a low percentage of normal forms in fertile men at first glance seems anomalous, bordering on an almost bizarre reduction *ad absurdum* for a fertile species, and contrasts with the much higher percentage of what are considered normal spermatozoa in semen from domestic species, biological reasons may well underly this difference. Unlike domestic species with multi-male mating systems, in which sperm competition within the female tract has driven the male to evolve to produce large numbers of good quality spermatozoa—a process often aided by selective breeding of domestic animals—our own species is serially monogamous and our spermatozoa pleiomorphic, since sperm competition rarely occurs. Thus, if fertile men are the correct reference group to produce reference values for semen, it is to be anticipated that only low percentages of normal spermatozoa will be found there. Only time will tell whether such a low reference limit for normal forms is clinically usefully.

Alternatively, the hint of quasi-theological fanaticism in the search for the rare, perfect sperm is hard to ignore. The best evidence that such “perfect” sperm are solely responsible for fertilisation is, however, at best correlative. Misplaced literalism here runs substantial risk of overinterpretation particularly as so-called deformed spermatozoa can fertilise oocytes in the laboratory thereby proving they are not totally devoid of fertilising ability as the overinterpretation implies. Interestingly, Menkveld [12] also argues that since the lower reference limit is so low, other attributes of spermatozoa, the abnormal forms, need to be examined, as discussed by Eliasson [4] and Auger [10]: does this constitute an eventual meeting of the minds?

Chris Ford [13], Rupert Amann [14], John Aitken [15] and Dolores Lamb [16] all address the concept that semen analysis should be more than descriptive and provide evidence of the pathology behind the semen picture. Ford stresses the poor prognostic power of semen analysis and desires a renaissance of diagnostic tests to elucidate why a man should need ICSI to provide the baby he could not produce naturally. The study of ROS measurement, DNA intactness and sperm proteins are mentioned as likely candidates to provide such valuable prognostic information; current lack of convincing evidence notwithstanding. Amann [14] considers that current assessment of sperm morphology semen analysis provides little information on testicular pathology and believes that tests should be designed to be interpretable as indicating whether damage to spermatozoa was done at spermiogenesis (indicating a testicular defect), within the excurrent duct system (a post-testicular defect) or at ejaculation (poor accessory gland function). He lists many sperm function tests, some of which are not, or could not easily be, performed, that are uncertain, inappropriate or appropriate for ascertaining defects at spermiogenesis in usual clinical settings. It is also doubtful whether the realistic clinical goals should be so focused on diagnosis rather than alleviation of infertility with the available means at hand. Lamb [16] also mentions tests that may not be useful but considers sperm-egg interaction tests (zona binding and hamster oocyte penetration test [or sperm penetration assay as the transatlanticists know it]) to provide additional worthwhile information, the latter even in the ICSI-era as a tool for training in the technique. She admits that, while DNA tests are not straightforward in interpretation, they can provide information potentially useful to evaluate an infertile male; the low uptake of these tests in clinical practice even by scientifically literate clinicians however suggests they vote, by the tests they order, against her proposition that these tests genuinely add diagnostic value. Nevertheless, which of the many sperm DNA tests is the most suitable is not answered and precisely how the data will be used is not outlined. For ICSI the spermatozoon injected cannot of course be studied, but if spermatozoa selected in a similar way (by morphology in a wet preparation) are collected they could be. ROS and antioxidants are mentioned as requiring more work but the detection of aneuploidy should be made a priority. Lamb [16] goes further to say that these tests are necessary, not only to provide evidence on the likelihood of treatment success, so that infertile couples can decide rationally whether to continue trying or come to closure and acceptance of the untreatability of their infertility, but also as they bring benefit to toxicology studies or clinical trials of drug testing. Aitken [15] states that although the widespread use of ICSI has killed off the development of tests that might predict the fertilising potential of spermatozoa, future tests should be designed to define molecular defects as part of far better diagnostic andrology. Again the uptake of such tests in clinical Andrology practice will be the impartial indicator of their practical utility.

Which reference values for human semen provided in WHO5 [7] are relevant brought much debate, as expected. Whether fertile men are a useful group for diagnosis is not only debatable but hotly debated. Michael Joffe's [17] article on what is meant by “normality” in relation to fertility, explains the difference between statistical normality (useful for epidemiology) and normality related to biological functions (useful for defining reproductive competence) and reveals the conundrum of an “abnormal” value being “normal”, a concept highly pertinent to reference values for semen. Joffe [17] considers that the WHO's recent fathers are suitable for male fertility assessment but Jens Bonde [18] argues that they are not relevant for epidemiological studies in which a comparable control is required by the study design. As noted previously, even these com-

plex arguments omit the issue of decision limits which may differ yet again depending on proven therapeutic goals (as for cholesterol). Ford [13] points out that even the concept of dichotomisation of continuous variables is incorrect, although used practically in virtually every other field of clinical chemistry. Joffe [17] concurs with Ford [13] and Aitken [15] that merely defining a value as subnormal neither identifies the pathology (useful for prescribing treatment) nor the proximate cause (useful for taking preventive measures). Boyd [11] also provides definitions of normality but takes the argument further, describing one- and two-sided reference intervals and the distinction between reference and decision limits. He also discusses the disadvantages of reference intervals and introduces several alternatives to them, considering that in the long-term multivariate likelihood ratios, a Bayesian concept embodying ongoing re-evaluation in the light of other relevant covariables, are more likely to emerge as providing a hint at a possible solution to the interpretation of data in cases where normal and abnormal distributions overlap.

The availability of the first ever valid reference values provided by WHO was addressed in several papers. Niels Skakkebaek's [19] presentation suggests that the new WHO reference limits for sperm number and percentage normal forms are too low claiming that the andrological management of patients may be affected by their use. He suggests that two sets of reference values are required, although when each is to be applied, and to which patient groups, is not made clear. The new "desirable" reference limits would be far higher than those in WHO5 [7] and based on data showing a linear relationship between sperm concentration (not sperm number) and chances of conception in one study which was included in the database from which WHO5 generated its reference values. His preferred cut-off would be the value below which rates of pregnancy rates start to fall in the study he favours ($40 \times 10^6 \text{ mL}^{-1}$). Of course, others have claimed the opposite, that the WHO5 reference limits would be far too high and involve overdiagnosis and overtreatment of healthy men labelled wrongly (harmful in itself) as subfertile. Clearly, the context of the data and application of a reference range makes a difference. For example, basketball coaches having trouble finding tall enough players for their teams may define their own ideal "reference range" for height, in effect a decision level for entry to the team, which is very different from the well established population norms. In this scenario, the decision (entry to team) is meaningful whereas the lack of effective treatments for oligozoospermic male infertility mean that setting decision limits at variance with population reference ranges is not justified by the decision it assists, in this case fostering only wasteful, ineffective testing.

Skakkebaek [19] is surprised that MacLeod, who once espoused that the vast majority of fertile men had sperm concentrations of $\sim 100 \times 10^6 \text{ mL}^{-1}$, should "change his view" to the lower limit being $20 \times 10^6 \text{ mL}^{-1}$. This may be nothing more than changing the description of data from the mean value to the 5th centile, now the accepted limit for pathology reference data. It may even be to MacLeod's credit that he gained greater insight into the issues of selection bias in data from his infertility laboratory [20].

A thoughtful perspective on population studies of semen analysis and its strengths and limitations is provided by Bonde [18], a pioneer in modern, rigorous large-scale studies of occupational and population based studies using semen analysis. He concludes definitely that there is no valid scientific basis to the alleged temporal or geographic decline in sperm outputs that have been so loudly proclaimed over the last two decades, noting that some corollary claims about environmental toxins are "... difficult to corroborate and almost impossible to refute..." rendering them into the realm of cult beliefs rather than testable science. He highlights that the reasons for the consistently higher sperm output of Finns compared with other Europeans remains entirely unexplained with even the non-environmental possibility of a genetic founder effects in Finland, the cradle of so much genetic epidemiology as one of the best studied genetic isolates [21], not being excluded.

If the long march of semen analysis to join other rigorously standardised chemical pathology tests is well underway, the even longer march of semen analysis to become a tool for clinical and population research outside the narrower realm of infertility practice, has barely begun. Even before the second large scale clinical study of semen analysis was published [22] after MacLeod's, there were already claims that sperm output was decreasing in the USA [23] and Denmark [24] although these claims were never considered convincing. Then came the apocalyptic claims of the controversial 1992 Carlsen meta-analysis [25], cited over 1 000 times already, and proclaiming alarming decreases in sperm counts over decades, despite its fatally flawed methodology. With mounting flamboyance and urgency, it was further alleged that this "crisis" was brought about by xenoestrogenic endocrine disruptors [26]. However, critical analysis of these publications, based on misinterpretation of semen analysis, exposes a plethora of flaws of methodological misuse, misinterpretation of data and ignoring salient bias as well as the implausibility of allegations that exposure to minute doses of chemical estrogens would have effects during pregnancy, a highly estrogenic milieu. The last nail in the coffin of this hypothesis was the evidence that even massive doses of prenatal estrogens had no deleterious effects on male reproductive function [27], consistent with a recent review that failed to find any evidence for such effects in man [28]. Interestingly the arc of progress is such that the latest review by the original proponents of this hypothesis did not contain the word "estrogen" [29]. While such fear-mongering may enliven journalists and frighten politicians, it falls far short of high quality evidence genuinely needed to advance knowledge as the cornerstone

of Andrology in the 21st century. Illustrating the dictum about the importance of the unimportant, the preoccupation of researchers with untestable global claims is one major reason why progress over the last decades has been so disappointing in advancing understanding of the undoubted, but still unrecognised, genuine environmental causes of male infertility. This is in stark contrast to the major progress made in scientific understanding of the genetic causes over the same period.

In his summary, Aitken [15] concludes that while basic semen analysis will stay in one form or the other, most likely in WHO5's format, it will be complemented with new tests. These could include the application of the "omics" (prote-, glyc-, lipid- and metabol-) to spermatozoa, but he cautions that the data obtained will be meaningless without knowledge of the functional defects of the sperm cell and advances in epigenetics.

A worrying aspect, not of the WHO manual itself but its use, was not addressed in the invited chapters. This is the deference with which it is held, especially with widespread reporting of "WHO standards" as a "gold standard" for all situations. Uncritical acceptance of WHO "reference" values has led to their being used for completely unsuitable purposes lacking specific evidence, e.g. taking the percentage of grade "a" spermatozoa to decide whether prepared spermatozoa should be used for IVF or ICSI. The data in WHO5 [7] and its predecessors have always referred to spermatozoa swimming in semen and not of prepared spermatozoa in medium. If cut-off velocities are required for decisions on ART procedures, data should be obtained specifically for that purpose, and not taken out of context from the WHO semen analysis manual.

Not surprisingly, when values are taken out of context and/or uncritically, irrelevant information may be produced. The new manual attempts to dispel this unhealthy attitude. First, it provides choices of tests that depend on the needs of the laboratory—do you need to measure very low sperm numbers accurately? If not, don't. Do you need to assess normal morphological forms only for ART? If not, examine abnormal forms as well. Do you need a quick screen of sperm antibodies on seminal spermatozoa or a more detailed examination of seminal plasma-free spermatozoa? Your choice. Quoting "WHO 2010" will no longer suffice as a method—the particular method (chamber, stain) used, has to be added.

In addition to the Perspectives, this Issue contains two original publications and a Commentary by Gordon Baker [30] on one of them. Lu *et al.*'s contribution [31], on the performance of semen analysis in China shows how few laboratories adhere to WHO guidelines, mirroring what was found by Brazil and Pacey in their contributions. Trevor Cooper & Barbara Hellenkemper [32] pursue the line of reasoning in the Foreword that only with adequate numbers can concentrations of any cell type in semen be precise. They show that even with the WHO5-recommended procedure for round cell estimation (unchanged from that in earlier editions) it is not possible to assess the consensus cut-off round cell concentration of 10^6 cells mL⁻¹ with a sampling error of 5% (required for estimates of sperm concentration). Even with the modification suggested (lower dilution, greater volume assessed), only a concentration of 1.3×10^6 mL⁻¹ can be assessed with this precision.

In this Special Issue, the authors have risen to the challenge provided in the Foreword. We allowed latitude in style towards gut reactions over strictly evidence-based concepts, in order to highlight several important concepts and debates, some remaining unresolved, about the meaning of normal values and how to interpret reference limits about which many have thought, or raised in meeting discussions, but never before put in print. These issues, with respect to male fertility and infertility in general, and the WHO manual in particular, receive a good airing and we hope that they will educate readers to approach the new WHO manual with a realistic and critical appreciation of the balance its recommendations represent.

References

- 1 Jequier AM. Semen analysis: a new manual and its application to the understanding of semen and its pathology. *Asian J Androl* 2010; 12: 11–3.
- 2 Brazil C. Practical semen analysis: from A to Z. *Asian J Androl* 2010; 12: 14–20.
- 3 Pacey AA. Quality assurance and quality control in laboratory andrology. *Asian J Androl* 2010; 12: 21–5.
- 4 Eliasson R. Semen analysis with regard to sperm number, sperm morphology and functional aspects. *Asian J Androl* 2010; 12: 26–32.
- 5 World Health Organization. WHO Laboratory Manual for the Examination of Human Semen and Sperm-cervical Mucus Interaction, 3rd edn. Cambridge, UK: Cambridge University Press; 1992.
- 6 World Health Organization. WHO Laboratory Manual for the Examination of Human Semen and Sperm-cervical Mucus Interaction, 4th edn. Cambridge, UK: Cambridge University Press; 1999.
- 7 World Health Organization. WHO Laboratory Manual for the Examination and Processing of Human Semen, 5th ed. Geneva: World Health Organization; 2010.
- 8 Björndahl L. The usefulness and significance of assessing rapidly progressive spermatozoa. *Asian J Androl* 2010; 12: 33–5.
- 9 Yeung CH, Cooper TG, Nieschlag E. A technique for standardization and quality control of subjective sperm motility assessments in semen analysis. *Fertil Steril* 1997; 67: 1156–8.
- 10 Auger J. Assessing human sperm morphology: top models, underdogs or biometrics? *Asian J Androl* 2010; 12: 36–46.
- 11 Boyd JC. Defining laboratory reference values and decision limits: populations, intervals, and interpretations. *Asian J Androl* 2010; 12: 83–90.
- 12 Menkveld R. Clinical significance of the low normal sperm morphology value as proposed in the fifth edition of the WHO Laboratory

- Manual for the Examination and Processing of Human Semen. *Asian J Androl* 2010; 12: 47–58.
- 13 Ford WC. Comments on the release of the 5th edition of the WHO Laboratory Manual for the Examination and Processing of Human Semen. *Asian J Androl* 2010; 12: 59–63.
 - 14 Amann RP. Tests to measure the quality of spermatozoa at spermiation. *Asian J Androl* 2010; 12: 71–8.
 - 15 Aitken RJ. Whither must spermatozoa wander? The future of laboratory seminology. *Asian J Androl* 2010; 12: 99–103.
 - 16 Lamb DJ. Semen analysis in 21st century medicine: the need for sperm function testing. *Asian J Androl* 2010; 12: 64–70.
 - 17 Joffe M. Semen quality analysis and the idea of normal fertility. *Asian J Androl* 2010; 12: 79–82.
 - 18 Bonde JP. Semen analysis from an epidemiologic perspective. *Asian J Androl* 2010; 12: 91–4.
 - 19 Skakkebaek NE. Normal reference ranges for semen quality and their relations to fecundity. *Asian J Androl* 2010; 12: 95–8.
 - 20 MacLeod J, Wang Y. Male fertility potential in terms of semen quality: a review of the past, a study of the present. *Fertil Steril* 1982; 37: 126.
 - 21 Peltonen L, Jalanko A, Varilo T. Molecular genetics of the Finnish disease heritage. *Hum Mol Genet* 1999; 8: 1913–23.
 - 22 Rehan NE, Sobrero AJ, Fertig JW. The semen of fertile men: statistical analysis of 1 300 men. *Fertil Steril* 1975; 26: 492–502.
 - 23 Nelson CM, Bunge RG. Semen analysis: evidence for changing parameters of male fertility potential. *Fertil Steril* 1974; 25: 503–7.
 - 24 Bostofte E, Serup J, Rebbe H. Has the fertility of Danish men declined through the years in terms of semen quality? A comparison of semen qualities between 1952 and 1972. *Int J Fertil* 1983; 28: 91–5.
 - 25 Carlsen E, Giwercman A, Keiding N, Skakkebaek NE. Evidence for decreasing quality of semen during past 50 years. *British Medical Journal* 1992; 305: 609–13.
 - 26 Andersson AM, Jørgensen N, Main KM, Toppari J, Rajpert-De Meyts E, *et al.* Adverse trends in male reproductive health: we may have reached a crucial 'tipping point'. *Int J Androl* 2008; 31: 74–80.
 - 27 Wilcox AJ, Baird DD, Weinberg CR, Hornsby PP, Herbst AL. Fertility in men exposed prenatally to diethylstilbestrol. *New England Journal of Medicine* 1995; 332: 1411–6.
 - 28 Storgaard L, Bonde JP, Olsen J. Male reproductive disorders in humans and prenatal indicators of estrogen exposure. A review of published epidemiological studies. *Reprod Toxicol* 2006; 21: 4–15.
 - 29 Sharpe RM, Skakkebaek NE. Testicular dysgenesis syndrome: mechanistic insights and potential new downstream effects. *Fertil Steril* 2008; 89 (2 Suppl): e33–8.
 - 30 Baker HW. Status of semen analysis—comment on “A survey on the status of semen analysis in 118 laboratories in China” by Jin-Chun Lu *et al.* in *Asian Journal of Andrology*. *Asian J Androl* 2010; 12: 115–7.
 - 31 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.
 - 32 Cooper TG, Hellenkemper B. Improving precision in the assessment of round cell numbers in human semen. *Asian J Androl* 2010; 12: 111–4.