

EDITORIAL

Falling sperm counts and global oestrogenic pollution: postscript

Trevor G Cooper¹ and David J Handelsman²

Asian Journal of Andrology (2013) 15, 208–211; doi:10.1038/aja.2013.17; published online 18 February 2013

This Special Issue provides reflections after 20 years, summarized in 12 papers, on the controversial issue of alleged decline in human sperm output due to global oestrogen pollution by industrial chemicals.

Virtanen *et al.*¹ offer no specific critique of the original papers but focus on male reproductive function in their native country because Finnish men allegedly have better sperm production than other Scandinavians. Despite this, from longitudinal studies of semen quality among European populations, they believe that young Finnish men had a declining sperm output between 1998 and 2006. They interpret this as Finland's following a global trend of impaired male reproductive health. As these multicentre studies adopt an uncontrolled population sampling methodology, the lack of evidence that the men included by participating centres truly represent the general male populations from which they were sampled means that inferences, whether relating to either time or place, remain unconvincing.

In a thoughtful, articulate contribution, Anawalt² concludes that the Carlsen meta-analysis³ was flawed, its conclusions wrong and the putative explanatory hypothesis therefore lacks foundation. Nevertheless, he argues that these misadventures served as sentinels for potential perils of widespread man-made chemical pollutants, and raised scientific and community awareness of risks to male reproductive health. He considers it likely that some new industrial chemicals ubiquitous in the environment are harmful to male human reproductive health, an area of scientific inquiry that was largely ignored before these articles appeared. In channelling Rachel Carson, however, he overlooks the consequences of her unbalanced pronouncements on the hazards of DDT. While DDT may have harmed certain wildlife (though this is disputed), there is no evidence that DDT ever harmed any humans. Indeed, the elimination of DDT from public health use for malaria prophylaxis had disastrous health and medical impacts involving the resurgence of malaria, causing millions of excess deaths and contributing to drug resistance, a public health disaster detailed elsewhere.⁴ These salient unintended consequences of well-meaning but tunnel-visioned activism are less well known than deserved. Anawalt's sneaking admiration for what may seem harmless 'crying wolf' raises the question whether we can expect to make progress through sound science by accepting uncritically false alarms, however implausible, if they make flamboyant enough claims that attract public clamour. The endless opportunities for activists or crusading journalists to employ the hollow echo chamber of the internet for propagation

of 'scares' among the general population are hardly a responsible voice of science. Can we justify shouting 'fire' in a crowded theatre on the grounds that it serves as an impromptu fire drill?

Fisch and Braun⁵ concur with Anawalt² that allegations of a worldwide decline in sperm output do not withstand scientific scrutiny. Summarising post-Carlsen studies, they find that most claim no changes, with a minority claiming decreases or other ambiguous findings. Rather than considering that the preponderance of refuting evidence puts the matter to rest, they suggest that further investigation of the geographical and temporal variations in semen quality claimed. They make the key point that studies involving semen analysis are virtually always non-representative of their source populations, owing to participation or recruitment bias resulting from the low willingness of otherwise healthy, non-infertile men to provide semen samples. Yet they endorse claims of supposed geographical differences between centres, which simply substitute a geographical fallacy of comparability for the previous fallacies of alleged temporal changes. The common issue is that biased study sampling does not allow extrapolation back to the source population. In scientific logic, the non-reproducibility of sperm output between centres, whether differing in time or place, most probably simply reflects differences in the biased recruitment at each centre arising from subtle differences in recruitment practices and efficiencies to overcome low participation rates. It would really be a failure to learn the lessons of this controversy if we simply substitute one fallacy for another.

Ford⁶ focuses on laboratory Andrology and concludes that further studies are needed, as the previous reports are marred by methodological flaws in the semen analysis methods used. After thorough exposition, he suggests that only validated semen analysis methodology should be included in future large cohort studies that might gather data on the gestation and upbringing of children, the medical and environmental exposure of the parents, and genetic evaluation of the children, parents and grandparents, and which should help to fill the gaps in our knowledge on the genetic and environmental factors that determine a man's sperm count.

Pastuszak and Lamb⁷ agree with Fisch and Braun⁵ that numerous post-Carlsen studies show no consistent changes in sperm output over time. They discuss the 1992 meta-analysis as well as the subsequent community discussion surrounding it, concluding that no definitive linkage has been established between environmental factors or geography and the alleged decline in sperm output over time. They also

¹Tuen Mun, NT, Hong Kong SAR, PR China and ²ANZAC Research Institute, Concord Hospital & University of Sydney, Sydney, New South Wales 2139, Australia
Guest Editors for this special issue: Dr TG Cooper (ctrevorg@gmail.com) and Prof. DJ Handelsman (djh@anzac.edu.au)

focus on the dissemination of their findings to the public and support the call for more responsible and cautious interactions with science journalists to cap the more flamboyant exploitation of scientific data for media sensationalism and activist propaganda.

Nieschlag and Lerchl⁸ focus on some problems with the Carlsen paper and agree with Ford⁶ and Pacey⁹ that systematic errors due to changes in semen analysis methodology over time could explain some of the changes observed. This reflects the fact that for longitudinal studies, systematic errors such as in changes in methodology are intractable, unlike random errors which can be at least partly overcome by larger sample sizes. With Anawalt,² they accept the silver lining that the original defective study has drawn attention and research funding to the field of Andrology. However, their analysis echoes that of Pastuszak and Lamb⁷ with Fisch and Braun⁵ in concluding that new data from the same group, covering a 15-year period, provide no basis for any belief in allegations of widespread decreases in sperm output. Instead, they suggest that standardized methodology will yield a sound basis for analysis of changes, if any, in semen parameters, fertility and infertility.

In a cogent and thorough exposition, Pacey⁹ examines the popular and scientific debates about a possible decline in semen quality from their basis on retrospective semen analysis data. He also concurs with other authors that such comparisons are unsound because of systematic errors reflecting the fact that methods of laboratory Andrology have changed considerably since the 1950s, with developments in training and competence, emphasis on Quality Assurance and Quality Control, and attempts at standardisation of technique. With Nieschlag and Lerchl,⁸ Pastuszak and Lamb⁷ and Fisch and Braun,⁵ he agrees that the only large prospective study carried out to date, overlooking the problem of uncontrolled sampling, shows no change in sperm concentration over 15 years. He argues that this supports the contention that when laboratory methods are adequately controlled, no secular change in sperm counts are observed.

Although Safe¹⁰ argues that while there is no support for the hypotheses of endocrine-disrupting compounds and fertility, he considers the controversies are no nearer resolution today. His pessimism arises from the correspondence, following publications by reproductive biologists with significant contributions to this field, containing evidence-based scepticism on the claimed effects of bisphenol A. As the susceptibility of hypotheses to data-driven refutation is what demarcates science from cult beliefs, it is a concern if work in this field would remain impervious to the usual rules of scientific logic. Given the turbulence of frequently poorly-substantiated claims about detrimental health effects of so-called 'endocrine disruptors', it is imperative that regulators remain motivated by firm scientific standards in distinguishing between chemicals that are really harmful to human health, and the many more that are the background noise of false alarms.

Te Velde and Bonde¹¹ accept that concern about adverse effects of environmental pollutants is both understandable and warranted, but consider the concerns expressed by members of the Science Policy Briefing on Male Reproductive Health to be excessive and inappropriate. They examine the problem from the viewpoint of male and female fecundity, which may be adversely modified by intrauterine and postnatal influences related to our changing society and environment, but argue that *in utero* exposure of male fetuses to endocrine disruptors to produce the testicular dysgenesis syndrome proposed by Carlsen *et al.* in 1992 is too limited. They consider that changes in couple fecundity may be caused by exposure to environmental pollutants of the female partner as well as the male, and that changes in

lifestyle (smoking, stress, obesity, sexual behaviour and delay of child-bearing) should also be taken into account when addressing the impact of environmental pollutants. The extreme alarm expressed by the Science Policy Briefing is based on some misconceptions, notably overlooking methodological flaws as well as studies that did not indicate any notable change in sperm count during the last 10–15 years, that no evidence points to a decline in couple fecundity during the last 30–40 years and that birth rates are on the increase in many EU countries. Their conclusions are supported by a systematic analysis of global trends in male fertility over the last two decades which also confirms no consistent changes in prevalence of infertility.¹²

In a measured, thoughtful and scholarly contribution, Wilcox and Bonde¹³ review the scientific attention to testicular toxicology which was revived by the two 1992 papers this Special Issue focuses on. They note first the limitations of the retrospective data and the criticisms of the former paper, and second that direct oestrogenic activity of environmental contaminants is too weak to damage human foetal testicular development. Nevertheless, the more general topic of 'endocrine disruption' has emerged as an influential topic in environmental toxicology and epidemiology. They point out the evidence for low level environmental chemical interference with hormone signalling, and the large numbers of scientific papers on environmental endocrine disruptors demonstrating possible, but so far unproven, detrimental consequences for human health.

In the additional papers in the Special Issue, Lerchl¹⁴ provides the details of the possible exposure of humans to radiofrequency electromagnetic fields from various sources (e.g., mobile and cordless phones, base stations, TV and radio transmitters, WiFi adapters). He concludes that there is no evidence to indicate that radiofrequency electromagnetic field exposure causes adverse health conditions such as cancer, sleep disorders, headaches, *etc.* Examination of the *in vitro* and *in vivo* experiments on animals and spermatozoa that claim to indicate adverse effects on male fertility reveals many problems of irradiation dose and improper experimental design. He reminds us that this field of research is notoriously difficult and that experiments should be planned and performed with experts in this area.

Using a lucid evolutionary focus, Aitken¹⁵ also accepts that the Carlsen analysis was incorrect and the oestrogen hypothesis that 'explains' it lacked basis. Drawing attention to the likely divergences between sperm output and male fertility, he raises concern over the population consequences of fertility treatments which, by facilitating technology-dependent fertility of infertile men, may gradually increase the prevalence of genetic male infertility. Yet his explicit focus on the very long timescale, centuries rather than decades, counteracts the taint of eugenics that such concerns traditionally evoke; in any case, the quantitative impact would likely be not only slow but very small. His paper reflects the Zeitgeist among some scientists of a pessimistic mentality suffused with a sense of impending doom from 'chemicals' as environmental 'toxicants'. Yet man-made chemicals, especially those manufactured efficiently at industrial scale, as required for them to become potential environmental 'contaminants', remain relatively primitive in structure compared with many highly complex, naturally occurring chemicals (e.g., complex alkaloids) to which humans have always been naturally exposed in the environment, notably in foods. Further, it is misleading to think that such industrial chemicals, though evolutionarily 'unfamiliar' to our detoxification systems, are incapable of being metabolized by it. Virtually, all xenobiotic chemicals, including modern drugs, are metabolized in humans, resulting in complex arrays of metabolites with either increased or more often decreased

potency. Some are toxic, but many others, perhaps a large majority, are harmless.

While it is hard to discern the origin of such global pessimism, the present world view of increasing numbers of medical scientists contrasts with that of the pride in the scientific and technological achievements of the later twentieth century, motivated then towards the idealistic goals of improving the world (conquering disease, reducing poverty, hunger, improving comfort and convenience of life, *etc.*). Now scientists are often timorously fearful of 'chemicals' and yearning for the (imaginary) golden age when everything in the environment was 'pristine' and people lived more naturally—for their 40-year life expectancy. It is indeed a curious paradox of modern life that at a time when life expectancy, a convenient and compact measure of health, is longer than ever in human history, public concerns about apparently pervasive and allegedly detrimental effects of industrial chemicals' damaging human health are greater than ever. Ironically, these neo-Malthusian preoccupations originate within wealthy countries rather than in developing countries, where life expectancy has yet to catch up. Perhaps such concerns are a product of comfortable affluence, a modern luxury hobby, rather like the manicured home lawns and other forms of pecuniary emulation which constituted what Veblen identified as the conspicuous consumption emblematic of the growing affluence of the emergent mercantile middle class of early twentieth Century America in their desire visibly to emulate the leisure class.¹⁶

FINAL REFLECTIONS

The reports here have indicated that one or more of three major misconceptions underlie the problems of the Carlsen and other papers that suggest a global decline in sperm counts and male fertility: the general acceptance (i) that the population of men studied in each paper is representative of the source population from which it was drawn in place and time; (ii) that the semen analyses are valid and precise; and (iii) that semen analysis is an acceptable surrogate variable for fertility. For each of these cases, the reviewers generally agree that (i) obtaining study populations that accurately reflect semen quality in the source population in time and place is effectively unattainable; (ii) changes in semen analysis methodology, the inconsistent adherence to standardized methods, and the absence of reliable quality control raises doubt about accepting historical results at face value; and (iii) the lack of a clear relationship between semen quality and male fertility invalidates semen quality as a suitable surrogate.

Virtually, all the authors agree that the Carlsen meta-analysis was seriously flawed and untenable in its conclusions on world-wide falling sperm output. This is in keeping with the extensive agricultural animal data showing no changes in sperm production despite exposure to the same environment (but without the participation bias of the human studies) over the same time period.¹⁷ Nevertheless, 20 years on, the methodological flaws in study sampling design remain unlearned lessons to some, as exemplified by a recent paper claiming once again falling sperm output, this time aggregating retrospective data from 126 infertility clinics within one country.¹⁸ With its illiterate title, and lacking any information on the semen analysis methods used, the paper is a throwback to pre-Carlsen period in using data from men attending infertility clinics. The authors assume that men with sterile wives are a representative population sample. However, even in health systems with minimal barriers to access, the uptake of fertility services and the degree of focus on male infertility and its increasing modern treatments are highly variable, so that their constancy between and within centres over time cannot be ignorable as the pivotal assumption requires.

However, several authors still consider that low-level environmental pollutants are possible harbingers of trouble with male reproductive health. As such, there is a general recognition that these issues can only be resolved through examination in well-designed, rigorously controlled large-scale prospective studies, preferably those that examine fecundity and with careful control of recruitment bias, to determine whether there is any decline in fecundity or its imperfect surrogates, sperm output and semen quality.

Probably the inescapable conclusion to this controversy was best summarized by the prescient comment over 30 years ago, a decade before Carlsen, that '...human (male) fecundity cannot be effectively monitored without semen studies of randomly selected samples of national populations and cross-national comparisons of similar type....such studies are long overdue...' ¹⁹ to which the pioneer of semen analysis John Macleod replied '...more studies, national and international, in human fertility are desirable. For obvious reasons, randomly selected semen samples are extremely difficult to obtain'.²⁰ Perhaps these last two decades have only embroidered on that theme but only by generating more heat rather than light.

It is safe to say that the grievous methodological errors of the Carlsen meta-analysis are unlikely to be repeated, or at least accepted again in any major medical journal. It is less clear whether the reasons for its failure are widely and fully understood. Although any physical or mental phenomena can be subjected to objective scientific analysis so long as valid data are available, it is hubris to assume that science can answer all scientific questions. A major constraint is that science is wholly data-dependent, and where valid data are absent so, unavoidably, will be any scientific analysis. As the legendary statistician, John Tukey reflected in his *Sunset Salvo* at the end of his career 'The combination of some data and an aching desire for an answer does not ensure that a reasonable answer can be extracted from a given body of data'.²¹

On the other hand, the hasty misattribution of causality to oestrogen pollution is less easily avoided. Refutation having been denied by some but not all^{22,23} proponents, the oestrogen pollution hypothesis has transformed into a new field of enquiry under the pejorative misnomer 'endocrine disruptors' as well as migrating to a variant hypothesis with equally grand and speculative scope ('testicular dysgenesis syndrome') now attributed to a different cause but with little new or definitive human data. This purports to link increased incidence of testicular cancer with greater incidence of cryptorchidism, hypospadias and reduced sperm output.²⁴ Although there is an unquestionable but small world-wide rise over decades in testicular cancer incidence (confirmed by reliable prospective cancer registries), the associations with other components of this putative syndrome are neither verifiable nor related to each, other than the well-known linkage of cryptorchidism with testicular cancer.²⁵ Cryptorchidism and hypospadias are non-lethal birth defects for which birth statistics are either not routinely recorded or, if registered at all, not examined by well standardized objective methods. Hence, retrospective data on these minor birth defects of comparable scope and reliability with prospective cancer registries over the twentieth century do not exist. Thus, although the gradual increase in testicular cancer is unquestionable, the other three components of this new 'syndrome' are factually dubious as well as lacking a linkage to increasing cancer. Whether this new syndrome ever proves justifiable, its naming may be an example of the unicorn

syndrome—the coining of a memorable name which gives the illusion of reality to even imaginary or non-existent entities.

- 1 Virtanen HE, Sadow S, Vierula M, Toppari J. Finland is following the trend-sperm quality in Finnish men. *Asian J Androl*; e-pub ahead of print 14 January 2013; doi:10.1038/aja.2012.142.
- 2 Anawalt BD. The silent spermatozoon: are man-made endocrine disruptors killing male fertility? *Asian J Androl*; e-pub ahead of print 21 January 2013; doi:10.1038/aja.2012.148.
- 3 Carlsen E, Giwercman A, Keiding N, Skakkebaek NE. Evidence for decreasing quality of semen during past 50 years. *BMJ* 1992; **305**: 609–13.
- 4 Robert D, Tren R, Bate R, Zambone J. *The Excellent Power: DDT's Political and Scientific History*. Indianapolis, IN: Dog Ear Publishing; 2010.
- 5 Fisch H, Braun SR. Trends in global semen parameter values. *Asian J Androl*; e-pub ahead of print 7 January 2013; doi:10.1038/aja.2012.143.
- 6 Ford W. Ignorance but not bliss: too little is known about the determinants of semen quality. *Asian J Androl*; e-pub ahead of print 14 January 2013; doi:10.1038/aja.2012.158.
- 7 Pastuszak AW, Lamb DJ. Counting your sperm before they fertilize: are sperm counts really declining? *Asian J Androl*; e-pub ahead of print 21 January 2013; doi:10.1038/aja.2012.105.
- 8 Nieschlag E, Lerchl A. Sperm crisis: what crisis? *Asian J Androl*; e-pub ahead of print 14 September 2012; doi:10.1038/aja.2012.90.
- 9 Pacey AA. Are sperm counts declining? Or did we just change our spectacles? *Asian J Androl*; e-pub ahead of print 28 January 2013; doi:10.1038/aja.2012.165.
- 10 Safe S. Endocrine disruptors and falling sperm counts: lessons learned or not! *Asian J Androl*; e-pub ahead of print 24 September 2012; doi:10.1038/aja.2012.87.
- 11 Te Velde ER, Bonde JP. Misconceptions about falling sperm counts and fertility in Europe. *Asian J Androl*; e-pub ahead of print 3 December 2012; doi:10.1038/aja.2012.122.
- 12 Mascarenhas MN, Flaxman SR, Boerma T, Vanderpoel S, Stevens GA. National, regional, and global trends in infertility prevalence since 1990: a systematic analysis of 277 health surveys. *PLoS Med* 2012; **9**: e1001356.
- 13 Wilcox AJ, Bonde JP. On environmental threats to male infertility. *Asian J Androl*; e-pub ahead of print 21 January 2013; doi:10.1038/aja.2012.153.
- 14 Lerchl A. Electromagnetic pollution: another risk factor for infertility, or a red herring? *Asian J Androl*; e-pub ahead of print 24 September 2012; doi:10.1038/aja.2012.104.
- 15 Aitken RJ. Falling sperm counts twenty years on: where are we now? *Asian J Androl*; e-pub ahead of print 28 January 2013; doi:10.1038/aja.2012.167.
- 16 Veblen T. *The Theory of the Leisure Class*. New York: Dover Thrift Editions; 1899.
- 17 Setchell BP. Sperm counts in semen of farm animals 1932–1995. *Int J Androl* 1997; **20**: 209–14.
- 18 Rolland M, Le Moal J, Wagner V, Royere D, de Mouzon J. Decline in semen concentration and morphology in a sample of 26 609 men close to general population between 1989 and 2005 in France. *Hum Reprod* 2013; **28**: 462–70.
- 19 Sloan DG. Male fertility potential. *Fertil Steril* 1982; **37**: 126.
- 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 Tukey JW. Sunset Salvo. *American Statistician* 1986; **40**: 72–6.
- 22 Sharpe RM. "Additional" effects of phthalate mixtures on fetal testosterone production. *Toxicol Sci* 2008; **105**: 1–4.
- 23 Sharpe RM. Is it time to end concerns over the estrogenic effects of bisphenol A? *Toxicol Sci* 2010; **114**: 1–4.
- 24 Sharpe RM, Skakkebaek NE. Testicular dysgenesis syndrome: mechanistic insights and potential new downstream effects. *Fertil Steril* 2008; **89**: e33–8.
- 25 Akre O, Richiardi L. Does a testicular dysgenesis syndrome exist? *Hum Reprod* 2009; **24**: 2053–60.