Counting your sperm before they fertilize: are sperm counts really declining?

Alexander W Pastuszak1,2 and Dolores J Lamb1,2,3

Asian Journal of Andrology (2013) 15, 179–183; doi:10.1038/aja.2012.105; published online 21 January 2013

A purported global decline in sperm counts has been a source of controversy since the early 1990s. Numerous studies performed since then, as well as reanalysis of the original data, show either no changes, or even increases, in sperm concentrations over time. In this review, we discuss the 1992 meta-analysis that initiated the continuing debate on whether sperm concentrations are declining, and the community discussion surrounding it over the past two decades. We also highlight studies evaluating sperm concentrations performed since the initial study that produced different results, and conclude that no definitive cause for a decrease in sperm counts has been established, and the effects of geography and environment on sperm counts are unclear.

Approximately 15% of couples are considered infertile, with a male factor implicated in up to 50% of cases, underscoring the essential contribution of the male in conception. However, only over the past two decades have we begun to understand the degree to which men contribute to a couple’s infertility, which historically was attributed to the female partner. One quarter or more of infertile men have an unknown cause for their infertility, and while recent work suggests that many of these etiologies are genetic, we cannot exclude the impact of environmental factors on the male reproductive system, which have been incompletely defined.

Work over recent decades has identified a potential decline in sperm concentrations, with at least one explanation being an increase or change in environmental pollutant levels. In 1992, a meta-analysis of sperm concentrations derived from 61 papers dating between 1938 and 1990 involving semen quality from 14,476 men found that both sperm concentrations and seminal volumes had dropped significantly during that period. This was followed in 1993 by a manuscript suggesting that this decrease in sperm concentrations was partly due to environmental contamination with estrogenic compounds. However, there was little evidence identifying estrogens in the geographical areas where sperm concentrations appeared to be declining, or other evidence directly linking estrogenic pollution with declines in male fertility. These works triggered a continuing discourse and numerous investigations to evaluate the verity of these claims. However, current data have failed to demonstrate a decrease in either sperm concentrations or male fertility.

Owing to the large amount of unclear and confounding data, it is difficult to determine whether male fertility rates are truly declining. Fertility assessment in both men and women is made more difficult by the wide range of sperm concentrations that can result in pregnancy. Indeed, sperm count alone cannot define the fertility potential of a couple. Nevertheless, from the authors’ viewpoint, the available data do not support a global decline in sperm concentrations or male fertility. We highlight the limitations of the original work claiming this decline, and show that current data, including data by the authors of the original work, support the argument that sperm concentrations have overall remained stable or even increased. Furthermore, we find that there are few rigorously performed studies evaluating sperm concentrations, highlighting a need for such studies in determining the status of male reproductive function and the environmental factors that may affect it. Finally, we provide our views on the importance of peer review and the role of scientists as intermediaries with the press, and thereby the lay population, and how to maintain this relationship while providing accurate data without unnecessary sensationalism.

DECLINING SPERM COUNTS ACCORDING TO CARLSEN: FACT OR FICTION?

When Carlsen et al. published their paper in the British Medical Journal in 1992, the evaluation and treatment of male infertility was still in its youth. Semen analysis had been available for many decades, but the precision and accuracy of the tests had changed significantly over the preceding 80 years with the development of improved sperm counting chambers, attempts at standardization of the assays and perhaps most importantly, the development of proficiency programs to allow interlaboratory comparisons of accuracy. The variability in semen analysis results makes it difficult to determine differences between studies over the large time frame analyzed, 1938–1990, and valid comparisons are not possible owing to the different analytic methods employed. There is reason to suspect that comparisons of semen quality performed from one decade to the next are invalid, despite the authors’ claim that differences in semen analysis techniques over time did not exist.

First, the authors claimed that since there was no trend in hematology results between the decades, despite changes in technology, comparisons between semen analyses should stand. However, the different viscosities of blood and semen make the use of white cell pipettes to dilute semen (routine until the 1960s) inappropriate as it introduces large (25%) errors in volume measurement,
invalidating the conclusion that semen analyses from different laboratories over time are equivalent. In Figure 2 of Carlsen et al., four of the five papers evaluated before 1960 showed mean sperm concentrations above 100 million ml$^{-1}$, in contrast to the majority of the remaining papers published after 1960, suggesting that a methodological bias was present.

Carlsen et al. used mean sperm concentrations to evaluate temporal trends. However, the progressive lowering of the World Health Organization (WHO) reference standard for a ‘normal’ sperm concentration could have unsuspectingly decreased observed mean sperm concentrations over time, given that ‘normal’ men were selected for each study, and therefore outliers with sperm concentrations outside of the ‘normal’ range would have been rejected. The ‘normal’ sperm concentration of 20 million ml$^{-1}$ was initially set in a landmark study of 1283 men by McLeod and Gold in 1951, and even in that study mean sperm concentrations were 107 million ml$^{-1}$. If sperm concentrations followed a normal distribution, a progressive decline in the reference standard would not account for a progressive decline in mean sperm concentrations, on the basis of mathematical modeling of data used by Carlsen paper. However, the distribution of sperm concentrations is not normal, making the use of means, and the decreasing reference standard as a selection criterion, serious limitations of the Carlsen study, particularly when the data are analyzed by Gaussian linear regression. The authors should have used median sperm concentrations and geometric or logarithmic transformation of their data, as previously discussed, or other forms of statistical analysis that would remove these limitations. In an independent reanalysis of the Carlsen et al. data, Olsen et al. demonstrated that the use of alternative statistical methods to linear regression, including quadratic, spline fit and stair-step analyses, the last of which showed a catastrophic decline in the 1960s, resulted in constant or slightly increased sperm concentrations over time. In all of these models, the data fits were as good as or better than the linear regression model used in the Carlsen paper.

Second, the authors produced a meta-analysis of all data as an aggregate, without subgroup analyses. This is a significant study limitation, particularly as the authors later attempted to explain the decline in sperm concentrations by implicating environmental estrogens as culprits. The data used in the Carlsen meta-analysis were derived from around the world, and one would expect that if environmental factors were responsible, these factors would vary with region and country. In fact, when excluding data from studies before 1960, owing to possible methodological biases, and examining only data from the United States, the only country from which multiple studies were analyzed from both the 1970s and 1980s, no significant change in median (interquartile range) sperm concentration is observed by decade (61.4 (56.4–81.1) million ml$^{-1}$ 1970s, 67.0 (61.5–78.0) million ml$^{-1}$ 1980s, P=0.83) (Table 1). In addition, when the data were evaluated by region and date, no differences in sperm concentrations were observed between decades within each geographic location. These analyses of the data evaluated by Carlsen et al. suggest no secular trend in sperm concentrations over at least two decades, and are in agreement with the findings of a 1979 study in which semen analyses performed in a single laboratory from 14 476 men from 1951 and 1966–1977 were examined. An independent analysis of the Carlsen data of the 48 studies published since 1970 demonstrated a significant increase in sperm concentrations, another finding that no change in sperm concentrations globally, with a regional decrease in the United States and other analyses found significant geographic variability of sperm concentrations, with decreases in the United States and Europe, highlighting the limitations of inappropriate statistical methods or the inherent limitations of the original data, including the relative importance of one semen variable over another.

The Carlsen et al. study was also limited by selection biases intrinsic to the 61 individual papers; the men studied may not have been truly representative of their underlying populations, given the potential for self-selection bias, and some men had proven fertility whereas others were unsolicited for fertility, and thus considered to be representative of the general population. This introduced potential differences in sperm concentrations within the population studied. The abstinence period before production of a semen sample, well known to affect sperm concentrations, was poorly controlled for, particularly in the larger studies whose subjects were prevasectomy men.

It is significant that there were few data during the first 30 years of the analysis, with <10% of the studies representing 40% of the time analyzed and containing only 184 (1.2%) of the total number of men evaluated. The earlier data in this study contained much higher mean sperm concentrations than the subsequent data, and analysis of the later data showed no further decline in sperm concentrations, for which no explanation has been given. If one were to implicate social trends, such as decreased sexual abstinence or migration to warmer climates, in a decline in sperm concentrations, these effects should have been noticeable before 1938 if these trends were ongoing at the time, as suggested by alternative statistical models of the Carlsen paper data.

Can there be a consensus on sperm counts?

We recognize that sperm count alone, except in men with azoospermia, cannot be predictive of fertility, as defective sperm function may impair the capacity to achieve fertilization in vivo. Indeed, the fertility potential of

---

**Table 1 Median sperm concentration (interquartile range) by global region by decade**

<table>
<thead>
<tr>
<th>Region</th>
<th>1970s</th>
<th>1980s</th>
<th>$P^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of studies</td>
<td>Number of men</td>
<td>Median sperm concentration (million ml$^{-1}$)</td>
</tr>
<tr>
<td>United States</td>
<td>7</td>
<td>1868</td>
<td>61.4 (56.4–81.1)</td>
</tr>
<tr>
<td>Europe</td>
<td>4</td>
<td>162</td>
<td>87.7 (83.0–99.9)</td>
</tr>
<tr>
<td>Middle East</td>
<td>0</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Africa</td>
<td>0</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>South America</td>
<td>1</td>
<td>185</td>
<td>67.6 (mean)</td>
</tr>
<tr>
<td>Asia</td>
<td>2</td>
<td>80</td>
<td>77.9 (64.6–91.1)</td>
</tr>
<tr>
<td>Australia</td>
<td>0</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

ND, no data; (mean), mean sperm concentration in instances when only a single study for a given global region was available.

$P$ value determined by using the median test.

---

**References**

the couple is important. In recognition of this caveat, numerous other studies have been designed to evaluate sperm concentrations as a relative indicator of ‘male fertility’ since the Carlsen et al. work was published. These studies differed in their conclusions, with some finding decreases in sperm concentrations and others an increase or no change (Table 2). Collectively, these studies have reported semen quality from men around the world, most have been retrospective, and comparisons between studies are difficult because of the different statistical approaches used. Many studies reported mean sperm concentrations, casting doubt on the validity of the comparisons, though most have shown no change in sperm concentrations over time.

In 1994, Ginsberg et al.19 reviewed sperm concentrations from partners of 260 women presenting for fertility treatment in London over a 6-year period, as a function of their water source. They found a significant decrease in mean sperm concentration and motility in the group living within the Thames Water Area, and suggested an effect of environmental factors on sperm concentrations. Notably, the authors found a parametric distribution of sperm concentration, and used the t-test for statistical comparisons. In 1995, Auger et al.20 published a study examining semen analysis data in 1351 fathers, 96% of them Caucasian, presenting to a single sperm bank in Paris from 1973 to 1992. The authors found an annual decline of 2.6% in sperm concentration. These authors observed a non-parametric data distribution and controlled for the abstinence period. In addition, semen analyses were performed by a single laboratory without major changes in technique or personnel, and a relatively homogeneous population further improved the study. However, even though men heralded from the Paris area, data were not analyzed by the neighborhood in which they resided, a potential limitation, given that environmental pollutants may vary locally. Of note, a 1996 study by Bujan et al.21 analyzing sperm concentrations of 302 sperm donor fathers between 1977 and 1992 from the Toulouse area of France found no changes in sperm concentration, after correction for donor age and abstinence period, a contrast with the Paris study seen as further implicating the environment around Paris as a cause of the decline there. Many other studies22–24 appeared during the same year, evaluating sperm concentrations in the United States and Finland, but also finding no changes in sperm concentrations over time with differences in mean sperm concentrations dependent on geographical location. In 1998, Younglai et al.25 evaluated sperm concentrations in 48 968 Canadian men across 11 fertility centers and found an annual 1.44% decrease in sperm concentration between 1984 and 1996. However, no change in sperm concentration was observed with time when analyzing data from each individual fertility center. A 1999 study26 examining semen characteristics of 1927 donors in Denmark from 1977 to 1995 showed a decline in mean sperm concentrations over time as well as seasonal variation.

The above studies demonstrate the challenge of reliably identifying differences in sperm concentrations between groups, and also highlight the potential contributions of geography and environment in this variability, the true effects of which remain to be elucidated. Others have focused on geographical and seasonal variations in sperm concentration. Two studies by Jørgensen et al. in 2001 and 200227–28 evaluated semen quality across Europe, finding significant differences in sperm concentration by country, as well as seasonal variations in sperm concentrations. Seasonal variations in sperm concentrations have also been identified in the United States, with highest values during the winter and lowest during the fall.29 In addition, a North-South gradient in sperm concentrations has been observed in men from the United States, with higher values in colder climates.30 In China, up to 63% of men have at least one semen parameter value lower than the WHO lower reference values (2010), with potential implications for fertility.31–33 Other groups have also looked at sperm concentrations as a function of birth date, finding lower sperm concentrations in men born more recently.34,35

Several recent studies have readdressed the question of whether there has been a secular trend towards a decline in sperm concentration. In 2011, Axelsson et al.36 examined semen from 295 Swedish men presenting for medical examination before military service between 2008 and 2010, and compared them with those from a similar cohort from 2000 to 2001, and found no differences in values of any semen parameter. The same year, Jørgensen et al.37 presented their prospective evaluation of 858 Finnish men presenting for routine examination at age 18–19 years over 8 years and found a decrease in semen quality. Non-parametric analyses and a prospective cohort were used, increasing the quality of the work. However, both of these studies were relatively small compared with some prior studies. Most recently, Jørgensen et al.38 published a cross-sectional, prospective study of semen quality from 4867 Danish men between 1996 and 2010 compared with that from earlier studies of men from the same geographical region. In this study, authored by one of the authors of the Carlsen meta-analysis, sperm concentrations and total sperm counts increased over time. This is an important study given its prospective nature, large sample size and the single geographical location from which subjects were derived. Nevertheless, while relatively strong conclusions can be drawn about semen

Table 2 Summary findings of studies evaluating sperm concentrations over time, published after 1992

<table>
<thead>
<tr>
<th>Paper</th>
<th>Paper year</th>
<th>Number of men</th>
<th>Change in sperm concentration</th>
<th>Time period</th>
<th>Years evaluated</th>
<th>Geographical area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auger et al.20</td>
<td>1995</td>
<td>1351</td>
<td>Decrease</td>
<td>19 years</td>
<td>1993–1992</td>
<td>France</td>
</tr>
<tr>
<td>Bujan et al.21</td>
<td>1996</td>
<td>302</td>
<td>No change</td>
<td>15 years</td>
<td>1977–1992</td>
<td>France</td>
</tr>
<tr>
<td>Fisch et al.24</td>
<td>1996</td>
<td>1283</td>
<td>Increase</td>
<td>25 years</td>
<td>1967–1994</td>
<td>United States</td>
</tr>
<tr>
<td>Paulsen et al.25</td>
<td>1996</td>
<td>510</td>
<td>No change</td>
<td>21 years</td>
<td>1972–1993</td>
<td>United States</td>
</tr>
<tr>
<td>Vierula et al.26</td>
<td>1996</td>
<td>5719</td>
<td>Decrease</td>
<td>27 years</td>
<td>1980–1995</td>
<td>Finland</td>
</tr>
<tr>
<td>Younglai et al.27</td>
<td>1998</td>
<td>48 968</td>
<td>Decrease</td>
<td>13 years</td>
<td>1984–1996</td>
<td>Canada</td>
</tr>
<tr>
<td>Handelsman et al.28</td>
<td>1997</td>
<td>689</td>
<td>No change</td>
<td>15 years</td>
<td>1980–1995</td>
<td>Australia</td>
</tr>
<tr>
<td>Gyllenborg et al.29</td>
<td>1999</td>
<td>1927</td>
<td>Decrease</td>
<td>18 years</td>
<td>1977–1995</td>
<td>Denmark</td>
</tr>
<tr>
<td>Carlsen et al.30</td>
<td>2005</td>
<td>158</td>
<td>No change</td>
<td>4 years</td>
<td>1997–2001</td>
<td>Denmark</td>
</tr>
<tr>
<td>Axelsson et al.31</td>
<td>2011</td>
<td>295</td>
<td>No change</td>
<td>10 years</td>
<td>2000–2010</td>
<td>Sweden</td>
</tr>
<tr>
<td>Jørgensen et al.32</td>
<td>2010</td>
<td>858</td>
<td>Decrease</td>
<td>8 years</td>
<td>1998–2006</td>
<td>Finland</td>
</tr>
<tr>
<td>Jørgensen et al.33</td>
<td>2012</td>
<td>4867</td>
<td>Increase</td>
<td>14 years</td>
<td>1996–2010</td>
<td>Denmark</td>
</tr>
</tbody>
</table>
quality of Danish men, the variability in semen quality observed elsewhere make the results of this study difficult to extrapolate to a continental or global scale.

Thus, the global consensus on changes in sperm concentrations indicates that there has been no change in this variable during the past century.

**SPERM AND THE ENVIRONMENT: ARE WE MISSING A TOXIC ELEPHANT IN THE ROOM?**

The authors of the original meta-analysis suggested that environmental estrogens are culprits in decreasing sperm concentrations across the globe. They cited work indicating that environmental estrogen levels have increased over the past several decades in drinking water, our diets and our environments. While it certainly is interesting that birth rates appear to have declined over the past half century in developed countries, implicating possible environmental factors, there is little evidence establishing a causal link between estrogens alone and a decline in semen parameter values in human males. In addition, this does not take into account the use of contraceptives or other social factors that may result in a decline in birth rates. In fact, any link between male fertility and many environmental factors is tenuous. 39

Many chemicals including glycol ethers, 40 organo-chlorines, 41 exogenous estrogens and others 42,43 have been implicated as factors in male infertility but mainly in animal models. In addition, there is a lack of databases cataloguing the effects of chemicals on male reproduction, making it difficult to assess the true effects of these chemicals, if any, on human male fertility. 44

High levels of estrogens in the male can result in reduced Leydig cell function and sperm counts in the adult should follow. Sperm production is decreased by environmental exposure to other toxins. While several toxins have been investigated in animal studies, they have not been used in human studies, and exposure levels in animal experiments are typically significantly higher than corresponding levels in humans (reviewed 45).

**OF SPERM COUNTS AND DATA ANALYSIS: LESSONS LEARNED FOR THE FUTURE**

The significant number of studies demonstrating variable semen quality, while frustrating to those attempting to distil the data, is very informative. The Carlsen meta-analysis led to detailed analysis of many other papers that support the observation that how one’s data are derived and analyzed is critical in determining outcomes; incorrect statistics will lead to incorrect conclusions, but these errors were not detected by referees who should have prevented publication of substandard studies.

Whereas assessing male fertility status from semen quality is one of the more challenging clinical questions, because of the baseline variability in semen quality even among normal fertile males, rigorous analysis of such data is imperative in obtaining reliable, repeatable conclusions. That the Carlsen et al. finding spurred a decades-long debate is encouraging, as it highlights the interest of the community in the subject and the power of the scientific community in addressing problems, and this should be fostered. Nevertheless, even the most recent results of the largest prospective study on semen quality are not definitive, not because of inappropriate selection of subjects or data analyses, but because of the inherent limitations of good research—a single geographical location for deriving a relatively small, homogeneous, subject pool. Thus, more rigorous work should be considered to define variations in semen quality as a function of the population selected and geographical location. Such endeavours may ultimately highlight the potential mechanisms by which variations in semen quality occur.

Another topic of interest to the public, as this topic is, 46 is how to bridge the gap between the scientist and the layperson. It is both easy and tempting to report the results of any research study as absolutes, with little consideration for the details of the methodology that may cast doubt on the findings. In fact, many lay press authors do this in an attempt to make such findings ‘black and white’ for their audiences. However, this can lead to an incomplete or misunderstanding of the work by non-scientists and misinformation for the general public. Thus, any discussion with journalist groups, or publication by a researcher acting as a journalist, should include criticisms of work that should be made clear in the resulting publication. Furthermore, journalists should be made more aware of the fact that all studies have limitations and should strive to highlight these limitations in their publications in order to increase public awareness of the fact that all scientific investigations have limitations. Data should be presented objectively, and if Stephen Hawking could present complex physics to general audiences in an unbiased manner and be understood, all of us should work towards such abilities.

The debate over sperm quality is not over, although the consensus based on current data indicates that sperm counts are not declining. Future work will doubtless shed light on further contributions of geography, environment and lifestyle factors that could contribute to such secular variations in sperm concentrations, and will help us truly understand the landscape of this problem.

**COMPETING FINANCIAL INTERESTS**

All authors declare that there are no competing financial interests.

**ACKNOWLEDGMENTS**

This work was supported in part by the AUA Foundation Russell Scott Jr, MD, Resident Research Award (AWP), as well as NIH grants P01HD36289 from the Eunice Kennedy Shriver National Institute for Child Health and Human Development, National Institutes of Health (DJI) and IR01DK078121 from the National Institute of Kidney and Digestive Diseases (DJI).

---

Are sperm counts really declining?
AW Pastuszak and DJ Lamb

Paulsen CA, Berman NG, Wang C. Data from men in greater Seattle area reveals no downward trend in semen quality: further evidence that deterioration of semen quality is not geographically uniform. Fertil Steril 1996; 65: 1015.