



# Risk in Perspective

## The Sperm-Count Debate



Lorenz Rhomberg, Ph.D.

*"The popular press has reported the debate as a see-saw battle, now favoring the proponents and now the debunkers, depending on the result of the latest study to appear."*



Sonia Hernández Díaz, M.D., M.P.H.

In 1992, a key paper by Elizabeth Carlsen and coworkers appeared in the *British Medical Journal*; it argued that analysis of published estimates of sperm counts in human males shows an alarming, worldwide decline of about 50% from the 1930's to the 1990's. This paper has spawned a flurry of activity, as scientists around the globe struggle to determine whether the apparent decline is real or simply an artifact of changing measurement methods or unrecognized biases in the way data are collected. The results to date have been mixed—many papers report further data showing no decline in sperm counts over time, while others tend to support the original hypothesis. Arguments and counterarguments regarding proper analysis and interpretation of historical sperm-count data have been traded. The popular press has reported the debate as a see-saw battle, now favoring the proponents and now the debunkers, depending on the results of the latest study to appear. As in many scientific debates, the introduction of more data and increased scrutiny of analytical methods has served to show that the problem is more complex—and the final conclusion more elusive—than originally appreciated. In this issue of **RISK IN PERSPECTIVE**, we examine some of the issues and the state of the evidence regarding the hypothesis of sperm-count decline in the human population.

### SPERM COUNT

Beginning at puberty, the testes of human males begin producing many

millions of sperm cells per day, with the number initially rising and then gradually declining with age. Sperm count is a measure of the number of sperm cells in one milliliter (about 1/30th of an ounce) of seminal fluid. Researchers have measured sperm counts in connection with fertility studies and in donors to sperm banks. (It is the published results of such studies from previous decades that Carlsen and colleagues analyzed.) Normal sperm counts range widely among men and among studies, from highs of well over 100 million sperm/milliliter down to 40 million or less. Fertility (the ability to conceive children) is rarely compromised until counts drop below 20 million, and even men with counts under 5 million may be partially fertile.

Measured sperm counts are subject to many variable factors, and a good deal of the controversy surrounding sperm count data has to do with whether particular studies have adequately controlled for them. Age has already been mentioned; if studies from recent decades contain a higher proportion of older men (whose counts have declined as they age), then a false time-trend could appear, reflecting the changing mix of subjects rather than a real decline in sperm production over time for men of a given age. Some analyses control for age of donor statistically, but not all studies have the requisite age information. Similarly, sperm count will be sharply affected by the duration of sexual abstinence before donation, and only a few studies have

the data to control for this effect. There is even an annual cycle in human sperm production, being higher in the winter months and lower in the summer. These and other such factors lead to a great deal of variation in sperm count, even for the same man measured on different days. Even in well conducted studies, the person-to-person variation is large. Moreover, the laboratory methods for preparing and counting samples have changed over the years, and earlier and later counts may not be strictly comparable. The concern is whether temporal trends (i.e., trends over the years) in these extraneous factors are leading to false indications of a time-trend in sperm count (which should instead be interpreted as a trend in the average age or abstinence time of sperm-count study subjects).

### SCRUTINY OF STUDIES

The Carlsen study is a so-called meta-analysis; it examines the sperm count reports from 61 different studies published over a span of several decades. While using all the available published data has benefits, it also means that studies using different methods, on different numbers of men, and from widely different geographic locations must all be presumed to be comparable. Critics of the Carlsen study have questioned the comparability. Discussion has focused on three areas: selection of subjects, statistical analysis, and the potential for confounding factors (such as those mentioned above) to generate false indications of time trends.

Regarding subject selection, some studies examined potential donors to sperm banks, others only accepted donors; some examined vasectomy candidates while others studied patients at fertility clinics that were judged to have "normal" sperm counts. Different selection criteria focus attention on different segments of the entire male population, and these different segments vary in their average sperm count. The med-

ical definition of a "normal" sperm count was revised in the 1970's (from 60 million to 40 million), and so some earlier studies may have excluded men with lower counts, artificially boosting the apparent average count. Statistical discussions have focused on how the many studies with very few subjects should be weighted relative to the more reliable large studies. Some critics have also pointed out that, instead of a continuous, gradual decline in counts over time, the data show that most studies before about 1970 have similarly high counts, while those after that date have lower ones, but that no trend (or perhaps even a slight rise) is apparent after 1970.

The potential role of confounding factors such as age, duration of abstinence, and others has been mentioned. In 1996, a key observation was made by Dr. Harry Fisch and colleagues in the journal *Fertility and Sterility*. Sperm counts tend to vary from geographic region to region, with most American studies (and in particular, New York studies) showing high counts, while many studies conducted in European and third-world nations show lower counts. Fisch and colleagues noted that many of the early studies in the Carlsen et al. trend analyses came from the U.S., and several large studies reported data from New York, while later studies in the Carlsen database were predominantly from Europe, Asia, and Africa. Thus, the apparent time trend may actually reflect changes over time in where the studies were conducted.

### LOCAL TRENDS

Most of the criticism of the Carlsen study (and other meta-analyses) arises from its combining the results of studies done on different populations from different areas using different methods. They do not demonstrate that the apparent time-trends in world-wide sperm count are spurious, but they do raise questions regarding whether what

is being observed is a genuine change or an artifact of the combination of studies. In recent years, a number of studies have been published that report on sperm-count data from single regions (and even single clinics) over time. Within such a setting, the methods of measurement, subject recruitment biases, and base population all remain relatively constant, reducing the potential for spurious results.

The results of these studies of local time trends in sperm count have been mixed. A 1995 study by Auger and colleagues in the *New England Journal of Medicine* reported a decline among Parisian men over the last 20 years, and similar results have been reported in other areas (Sweden, Scotland, Belgium). In contrast, no declines (or even very slight increases) have been reported for other regions (elsewhere in France, Finland, Norway, and several U.S. sites such as New York, Seattle, Los Angeles, and Minnesota). The debate continues about the potential role of confounding factors, since some of the studies showing trends do not have the data to allow for changes in age or abstinence time among their study subjects over the decades. (And some studies that do have data on these factors demonstrate that they have indeed changed over time; it has been shown that neglecting to account for these changes would generate a false appearance of a sperm-count trend.)

On balance, the original hypothesis of a world-wide, ongoing decline in sperm count has been much undermined by demonstration of how easily false trends can appear in such data and by the illumination of the role of regional variation in sperm counts. Some more rigorous local studies demonstrate declines, however, while others show no such trends. The causes and impact of the geographic differences in sperm count remain unexplained. What is clear is that there is no consistent evidence of a world-wide effect on human

male fertility over the decades. While the "battle of the studies" reported in the press has cooled down, a period of sober, careful scientific study remains before the sperm-count phenomena are completely elucidated.

## MECHANISMS AND CAUSES

Among the reasons that the sperm-count question has received so much attention is that some scientists have speculated that the cause for a widespread decline may be the increasing exposure to chemicals in the environment, especially so-called endocrine disrupting agents that may act by mimicking the body's own hormones, perhaps affecting reproductive and developmental patterns that are under hormonal influence. In the laboratory, some endocrine disrupting agents have been shown to lower the sperm counts of rats. In seeking potential environmental culprits, it is important to remember that what seems to be at issue is exposure during fetal development. Thus, any effect on sperm counts measured in today's 20-40-year-old males should be sought among exposures to their mothers at least 20-40 years ago or more. Effects on males assessed in 1970

(where the Carlsen study first shows an appreciable apparent effect) should be sought in exposure patterns of women in the 1930's and 1940's. While chemical use has changed over the decades, one should not ignore the potential effects of the destruction of war, economic upheaval, and changes in the sources and nutritional quality of food over this timespan, especially in view of the apparent differences in Europe, the U.S., and the third world in more recent sperm count averages.

It is important to note that there are no data on any particular chemical or even on general chemical exposure that are tied to the widespread general patterns of sperm count discussed above. High enough doses of some particular chemicals can affect human sperm count.

**Harvard Center for  
Risk Analysis**  
Harvard School of  
Public Health  
718 Huntington Avenue  
Boston, Massachusetts  
02115-5924

617 432-4497  
www.hsph.harvard.edu/  
organizations/hcra/hcra.  
html



100% recycled paper,  
all post-consumer fiber.

#### FURTHER READING

Carlsen, E., Giwercman, A., Keiding, N., and Skakkebaek, N. 1992. Evidence for decreasing quality of semen during past 50 years. *British Medical Journal* 305:609-613.

Fisch, H., Goluboff, E.T., Olson, J.H., Feldshuh, J., Broder, S.J., and Barad, D.H. 1996. Semen analyses in 1,283 men from the United States over a 25-year period: no decline in quality. *Fertility and Sterility* 65(5):1009-1014.

Cooper, R.L. and Kavlock, R.J. 1997. Endocrine disruptors and reproductive development: a weight-of-evidence overview. *Journal of Endocrinology* 152:159-166.

#### PEER REVIEWERS:

George Gray, Ph.D.  
Sonja Ramstrom, M.S.

The best data on an endocrine-disrupting chemical, the former drug diethylstilbestrol (DES), which was given to some pregnant women in the 1950's and 1960's, shows no effect on sperm count in men exposed (as fetuses in the womb) to quite high levels.

#### QUESTIONS FOR THE FUTURE

In sum, the question of whether human sperm counts are changing over time has no clear answer in the mixed body of evidence now available. If there is an effect, it is certainly less universal and inexorable than was first suggested. What may be causing any such phenomenon is much less clear, and the hypothesis of chemical exposure is at present mostly speculative.

Sperm count is not the only measure of male reproductive health to be receiving attention. There are time trends in the incidence of undescended testicles, genital malformations, and testicular cancer that are also being closely studied. (We discussed this topic in the May 1997 *Risk in Perspective*.)

These questions are important enough that they deserve ongoing scrutiny. But in doing so, it is important to bear in mind the inherent difficulties in seeking historical patterns among studies conducted before today's knowledge of the potential pitfalls and confounding effects was established. The causes of geographic patterns in average sperm count remain puzzling, but seeking explanations may provide investigators with a window into associated factors and potential causes.

