The Question of Declining Sperm Density Revisited: An Analysis of 101 Studies Published 1934–1996

Shanna H. Swan,[1] Eric P. Elkin,[2] and Laura Fenster[3]

[1]Family and Community Medicine, University of Missouri, Columbia, Missouri, USA; [2]California Public Health Institute, Berkeley, California, USA; [3]Reproductive Epidemiology Section, California Department of Health Services, Emeryville, California, USA

In 1992 Carlsen et al. reported a significant global decline in sperm density between 1938 and 1990 [Evidence for Decreasing Quality of Semen during Last 50 Years. Br Med J 305:609–613 (1992)]. We subsequently published a reanalysis of the studies included by Carlsen et al. [Swan et al. Have Sperm Densities Declined? A Reanalysis of Global Trend Data. Environ Health Perspect 105:1228–1232 (1997)]. In that analysis we found significant declines in sperm density in the United States and Europe/Australia after controlling for abstinence time, age, percent of men with proven fertility, and specimen collection method. The declines in sperm density in the United States (approximately 1.5%/year) and Europe/Australia (approximately 3%/year) were somewhat greater than the average decline reported by Carlsen et al. (approximately 1%/year). However, we found no decline in sperm density in non-Western countries, for which data were very limited. In the current study, we used similar methods to analyze an expanded set of studies. We added 47 English language studies published in 1934–1996 to those we had analyzed previously. The average decline in sperm count was virtually unchanged from that reported previously by Carlsen et al. (slope = −0.94 vs. −0.93). The slopes in the three geographic groupings were also similar to those we reported earlier. In North America, the slope was somewhat less than the slope we had found for the United States (slope = −0.80; 95% confidence interval (CI), −1.37–−0.24). Similarly, the decline in Europe (slope = −2.35; CI, −3.66–−1.05) was somewhat less than reported previously. As before, studies from other countries showed no trend (slope = −0.21; CI, −2.30–1.88). These results are consistent with those of Carlsen et al. and our previous results, suggesting that the reported trends are not dependent on the particular studies included by Carlsen et al. and that the observed trends previously reported for 1938–1990 are also seen in data from 1934–1996. *Key words*: epidemiology, geography, regression analysis, semen quality, sperm density, trend. *Environ Health Perspect* 108:961–966 (2000). [Online 5 September 2000]
http://ehpnet1.niehs.nih.gov/docs/2000/108p961-966swan/abstract.html

In 1992 Carlsen et al. (*1*) stated that

> ...reports published worldwide indicate clearly that sperm density has declined appreciably during 1938–1990.

Subsequently, this conclusion has been supported by findings from some studies (*2–4*), but not by others (*5–7*). The critical issues raised concerning this study fall, broadly, into three categories. Some authors suggested that poor or highly variable data invalidated any inference about trends in sperm counts (*8,9*). Others questioned the validity of the statistical methods used in this analysis (*8,10,11*). Bias due to changing study populations (*12*) or confounding by factors such as age and abstinence time (time between sample collection and last ejaculation) were also suggested (*4,8*).

We conducted several analyses designed to examine these concerns. The first, published in 1997, reanalyzed the studies used by Carlsen et al. (*1*) to examine model selection, confounding, and selection bias (*13*). In that paper, we noted that estimates of mean sperm density from the United States and Europe declined somewhat more rapidly than had been reported by Carlsen et al. (*1*). In other parts of the world, where studies were few and

most were quite recent, there was insufficient data to evaluate this question. We also found that controlling for confounding bias, to the extent possible, provided additional support for the conclusions of Carlsen et al. (*1*) rather than reducing the estimated decline in sperm density. In the second analysis, published in 1999, we looked at sperm counting methods and the reliability of measurements from these historical studies (*14*). We found no evidence that counting methods had changed appreciably or that counts from older studies were less reliable than those from recent studies.

The current study extends our previous analyses in three ways. First, we conducted an independent literature review to evaluate possible bias in the selection of studies used by Carlsen et al. (*1*). Second, we examined the robustness of the models utilized in that analysis (and ours) by applying these models to an expanded data set. Finally, we assessed the consistency of post-1990 data with trends in sperm density from studies published before 1990.

Methods

Analysis of Carlsen et al. study. Carlsen et al. (*1*) screened studies published from 1930 to

mid-1990 to identify studies that included estimates of sperm density. They excluded studies that included men in infertile couples, men who were referred because of genital abnormalities, and studies that selected men on the basis of their sperm count. Studies that used nonmanual methods for counting sperm were also excluded. Carlsen et al. (*1*) included 61 studies published between 1938 and 1990. The authors estimated the rate of change in mean sperm density as a function of publication year by fitting a simple regression model.

Current analysis. The current analysis includes 54 of the 61 studies analyzed by Carlsen et al. (*1*). As in our previous paper (*13*), we excluded three non-English language studies (*15–17*) because it was not practical for us to systematically review the non-English language literature on this subject. We also excluded two studies that included men who conceived only after an infertility work up (*18,19*), studies that did not meet the eligibility criteria of Carlsen et al. (*1*). Finally, we did not include any studies with less than 10 subjects, which resulted in two additional exclusions (*20,21*). The most recent study in Carlsen et al.'s analysis (*1*) and our 1997 reanalysis (*1*) was published in June 1990. To extend the study period, we conducted a search of Medline (National Library of Medicine, Bethesda, MD) for English-language studies published between 1990 and 1996 and found 19 that met these eligibility criteria. We also conducted a less systematic search of the 60-year period 1930–1990 and identified 28 additional eligible studies. Therefore, the current analysis is based on 101 English-language studies published in 1934–1996 (54 "Carlsen" studies and 47 "non-Carlsen" studies), each with at least 10 men and all satisfying the eligibility criteria published by Carlsen et al. (*1*). The 47 "non-Carlsen" studies are summarized in the Appendix.

Each of these 101 studies was reviewed independently by two of us to systematically

abstract detailed information on potential confounders and several measures of semen quality. These variables included mean (or median) sperm density, publication year, study location (state and country), study goal (to estimate population parameters, other), criteria for recruiting study subjects (proven fertility, prevasectomy, potential sperm donor, other), percent of men with proven fertility, semen collection method (masturbation into container, other, unspecified), sperm counting methods (manual, not reported), number of samples per individual, age (mean or range), and abstinence time (mean or range, protocol requirement if applicable). Information on the completeness of this information was also recorded.

Previous analyses, including ours (13), have looked at the trend in sperm density as a function of publication year. However, because time of sample collection always predated publication, often by several years, we decided to use the time of sample collection, or its estimate, rather than the year of study publication. For the 22 studies that reported the beginning and end of the sample collection period, which often spanned several years, we used the midpoint to estimate the year of sample collection. The median lag time from the midpoint year to publication was 3 years for these studies. Therefore, when the dates of sample collection were unavailable, we subtracted 3 years from the publication year to estimate the year of sample collection. Finally, to obtain intercepts that were more easily interpretable and to aid in convergence of more complex models, we subtracted 1,900 from the estimated year of sample collection.

The arithmetic mean sperm density was reported in all but six studies. For these six studies we estimated the difference between the arithmetic mean and the reported summary measure (median or geometric mean) using data from studies for which multiple summary measures were available. For the five studies that reported median sperm density only, we estimated the arithmetic mean by adding 12.0 to the median, whereas for the single study that reported only a geometric mean, we added 22.7 to approximate the arithmetic mean.

We followed an analysis strategy similar to the one we used previously (13). After conducting a simple linear regression, we stratified the 101 studies into three broad geographic groupings: North America (44 studies, published 1934–1996), Europe (34 studies, published 1949–1996), and other countries (23 studies, published 1978–1995). We then used multiple regression models (using procedures for linear and nonlinear regression as well as generalized linear models) to fit linear, step, spline, and quadratic models

(22). In these models we included confounders that were related to sperm density and/or year in univariate analyses. Interactions between year and region, which can indicate geographic differences in the rates at which sperm density changed, were examined in all multiple regression models. To assess the extent to which each variable confounded the relationship between sperm density and year, we calculated the slope (in the model without interaction terms) with and without that variable included in the model. The magnitude of confounding is estimated by the degree of discrepancy between these two estimates (23). As with previous analyses, data from each study were weighted by the number of men included in that study, and sperm densities are given in units of 10^6/mL.

Results

The estimated year of sample collection in these 101 studies ranged from 1931 to 1994 (publication year 1934–1996). As shown in Table 1, the majority of new studies were published after 1980. Mean sperm density and mean publication year from studies with and without information about year(s) of sample collection did not differ appreciably. The geographic distribution of these 101 studies, representing 28 countries and 19 states within the United States, was similar to that in previous analyses, but with a somewhat greater proportion of European studies (Table 2). We made two changes in

our geographic strata; the stratum we previously labeled United States is now denoted as North America in order to include a (new) Canadian study. In addition, Australia, which was previously included with European studies, is now included with "other countries."

Simple linear model. For comparison with Carlsen et al. (1), we first replicated their simple linear regression. As shown in Table 3, the slope for the regression line in the expanded data set (-0.94×10^6 mL/year; $p < 0.0001$) is very similar to that found for the original 61 studies (-0.93×10^6 mL/year; $p < 0.0001$). These estimates differ only slightly from the slope we reported in our 1997 analysis (13): (-0.95×10^6 mL/year; $p < 0.0001$). The fit of the regression line to the 101 data points is shown in Figure 1.

Assessing confounding and interaction. To select variables for our analysis, we initially included all variables for which we had abstracted data and we noted the percent change in the slope that resulted when we removed them one at a time. Several of these were unrelated to sperm density or publication year (change < 10%) and were dropped from further analysis. These variables were the number of samples per subject, whether the years of sample collection were reported, whether the arithmetic mean was reported, and purpose of the study. Although removing age changed the slope by only 1.2%, we included this variable in the final model because it is a basic demographic variable

Table 1. Publication year of studies in three analyses.

Publication year	Carlsen et al. (1)	Swan et al. (13)	New studies in current analysis	All studies in current analysis
1930–1959	10 (16%)	8 (14%)	2 (4%)	10 (10%)
1960–1979	17 (28%)[a]	14 (25%)[a]	3 (6%)	16 (16%)
1980–1989	33 (54%)[a]	33 (59%)[a]	23 (49%)	55 (54%)
1990–1996	1 (2%)	1 (2%)	19 (40%)	20 (20%)
Total	61	56	47	101

[a]Includes one study with < 10 subjects that was excluded from the current analysis.

Table 2. Geographic distribution of studies in three analyses.

Region	Carlsen et al. (1)	Swan et al. (13)	New studies in current analysis	All studies in current analysis
North America	28 (46%)[a]	27 (48%)[a]	18 (38%)	43 (43%)
Europe	17 (28%)[b]	15 (28%)[b]	20 (43%)	35 (35%)
Other	16 (26%)	14 (25%)	9 (19%)	23 (23%)
Total	61	56	47	101

[a]Includes two studies with < 10 men that were excluded from the current analysis.
[b]Includes one Australian study included in "other" in the current analysis.

Table 3. Results of fitting a simple linear regression model in three analyses.

Factor	Carlsen et al. (1)	Subset of Carlsen in current analysis[a]	All studies in current analysis[a]
Number	61	54	101
Publication years	1938–1990	1938–1990	1934–1996
Slope	−0.93	−0.95	−0.94
p-Value	< 0.0001	< 0.0001	< 0.0001
R^2	0.36	0.36	0.22

[a]Excludes non-English language studies and those with < 10 men.

often included in analyses of sperm density. The method of counting sperm was also included (although removing it changed the slope by only 6%). In this expanded set of studies, recruiting criteria and the percent of men with proven fertility were highly correlated, so only one of these variables (fertility) was retained for further analyses. The following variables were included in all subsequent multiple regression models: geographic region, age, abstinence time, percentage of men with proven fertility, method of counting sperm, and method of sample collection (Table 4). Of these, all but the method of counting sperm had been included in our previous analysis (*13*). Because one of the goals of this study was to examine the effect of adding new studies, we also kept a variable that indicated whether the study had been included by Carlsen et al. (*1*), even though removing it from the model had little effect on the slope. Despite the incompleteness of data on many covariates, the inclusion of the variables contained in Table 4 did improve model fit. When the simple linear model was compared with the multivariate linear model, including these covariates, the adjusted R^2 increased from 0.22 to 0.59.

In addition to including these covariates singly, we examined interaction terms to allow for different slopes in the three
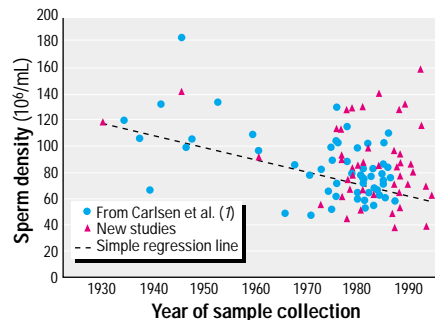

Figure 1. Mean sperm density in 101 studies published 1934–1996 and simple regression line.
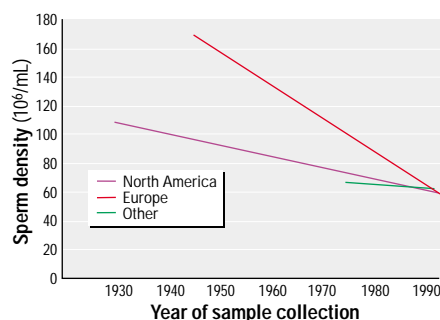

Figure 2. Interactive regression model for mean sperm density by year and geographic region, after controlling for proven fertility, abstinence time, age, specimen collection method, method of counting sperm, whether the study was included by Carlsen et al. (*1*), and interaction of region and study year.

geographic regions. In our previous analysis (*13*), the three slopes that we estimated differed considerably (−1.50, −3.13, and +1.56, respectively, for the United States, Europe/Australia, and other countries). In the current analysis the European slope (−2.35) still differed from that for North American (difference in slopes −1.55; CI, −2.90−−0.21), indicating significant interaction (Figure 2). Although we did include the slope of the best fitting line for other countries (−0.60), the fit to a linear model for data from these countries was not good and the confidence interval was very broad. Given the limited data, there was no evidence that this slope differed appreciably from those from other regions (Table 5).

Nonlinear models. We also fit a number of nonlinear models (quadratic, spline, and step) using the same set of covariates that were used for the linear model (Table 4). Olsen et al. (*8*) suggested that these models were preferable to Carlsen's simple linear model (*1*). In our 1997 analysis (*13*), we showed this was not the case, once geographic region and the interaction of region and year were included in the model.

In our previous study (*13*), we had not seen any difference between the spline and linear models except a slight (nonsignificant) change in the United States post-1970 (from −1.52 to −1.47; *p* for spline term = 0.97). When a spline model was fit to the current expanded data set, the pre-1970 North

American studies showed a somewhat steeper decline than those published after 1970 (−0.93 vs. −0.55), although this difference was still not significant (*p* for spline term = 0.71).

In our 1997 analysis (*13*), quadratic terms could not be estimated and we found no evidence of curvature within any of the three regions studied. In the present analysis, it was possible to estimate the quadratic term, but its addition did not improve the fit of the model; the quadratic terms were negligible and none approached statistical significance. Thus, we again found no evidence to support either curvature, or a "leveling off" in the rate of decline in recent years.

In our 1997 analysis (*13*), we also fit a step function and found a significant post-1970 decrease in sperm density in all regions relative to pre-1970 data (which was entirely from the United States). Again, results were similar in the current analysis. When a step function was fit, comparing the mean sperm density for North America before and after 1970, a large step was seen (138×10^6/mL vs. 113×10^6/mL; *p* for difference < 0.001). The pre-1970 mean from North America was also significantly higher than the mean for studies from other (*p* < 0.001), whereas the mean for all post-1970 European studies fell between the pre- and post-1970 North American mean.

Overall, the data fit these multiple regression models approximately equally (all

Table 4. Distribution of covariates retained in multiple regression models.[a]

Variable	No. of studies (*n* = 101)	Mean sperm density[b] (10^6/mL)	Mean year of sample collection[b]
Region			
North America	44	78	1970
Europe	34	87	1982
Other	23	65	1984
Included by Carlsen et al. (*1*)			
Yes	54	77	1974
No	47	68	1985
Age			
All men ≤ 40 years of age	23	97	1970
Some men ≥ 40 years of age	53	71	1980
No information	25	88	1963
Abstinence time			
Data reported: none < 3 days	14	81	1976
Data reported: some < 3 days	14	78	1983
No data reported: protocol restrictions reported	49	80	1977
No information	24	68	1976
Proven fertility			
Wife pregnant or post-partum	20	82	1968
At least 90% proven fertility (past)	31	67	1978
< 90% proven fertility (past)	8	84	1979
No information	42	86	1984
Method of semen collection			
Masturbation into container	70	70	1979
Other or no information	31	95	1971
Method of counting sperm			
Manual	66	71	1982
No information	35	79	1971

[a]Does not include geographic region, which is shown in Table 2. [b]Univariate/unadjusted, weighted by the number of men in each study.

adjusted R^2s were between 0.56 and 0.61), but not quite as well as the models fit the data in our previous analyses (13) (Table 5). As in our 1997 analysis (13), when multiple regression models that include terms for the interaction of geographic region and year are used, there is no support for the use of a nonlinear model.

Discussion

As we stated previously (13), control for confounding in these analyses can be only partial because of incomplete data. Therefore, it is possible that residual confounding remains. How large is this likely to be?

One of the strongest confounders in this analysis was the type of population studied. We examined this factor in two ways: the percent of men with proven fertility and the type of study population (sperm donor, pre-vasectomy, etc.) Because these variables were highly correlated, we retained only one (the percent of men with proven fertility) in the final model. When this variable was added to the other variables in the multiple regression model, it increased the magnitude of the slope considerably (37.2%).

Zavos and Goodpasture (24) reported that sperm concentration is higher when semen samples are obtained using a collection device during intercourse than when the same subjects collect samples by masturbation ($p < 0.01$), a result that has been reported by others (25). In the current analysis, studies that did not require collection by masturbation tended to be earlier (mean publication year 1970 vs. 1978). Therefore, this variable was a strong (positive) confounder; when it was added to the model, the magnitude of the slope decreased 34.1%.

Carlsen et al. (1) required that sperm be counted by manual methods in all the studies that they included in their analysis. Nevertheless, because manual counting devices have changed somewhat over the study period, when reviewing these studies, we abstracted information on the specific counting method that was used. When the particular counting device was not specified, we assumed it was manual. Nonmanual methods are a relatively recent advance and are still considered experimental, so that studies that use nonmanual methods are likely to specify the use of such methods. In 62 of these 101 studies, the counting device was specified to be the hemocytometer, the method that has been continually recommended by the World Health Organization since 1980 (26,27). The only other counting method that was specified, the Makler chamber (28), was mentioned in only 2 studies of 101 studies. Thus, we found no evidence that the introduction of newer counting devices has resulted in lower sperm counts.

In fact, when systematic changes have been introduced by newer methods, they tended to result in higher counts (14). In any case, this variable appeared to have little effect on the observed decline in sperm density.

Some researchers have criticized the use of sperm count estimates from early in the study period, arguing that greater measurement error was likely in these historical studies. Greater imprecision in earlier studies could not have produced the negative slope we observed in Western countries. A change in the variability of sperm counts would, however, violate a basic assumption underlying the regression methods used in these analyses, the assumption of constant variance. Was this assumption justified? To answer this question we looked for a trend in the standard deviation of sperm density in these historical studies. We modeled the standard deviation (which was reported in 34 studies) as a function of year and found no evidence of a trend (slope = −0.24; p = 0.22) (14). We also used a multiple regression model to examine possible confounding of this relationship, but found no evidence of this. We concluded, therefore, that there has been no significant change in the standard deviation of sperm density over time.

Geographic region and the interactions of region and year were important covariates in these analyses. However, these geographic groupings are large and heterogeneous. For example, the category "other countries" included Thailand, India, Hong Kong, Brazil, Australia, Kuwait, Nigeria, Israel, Libya, Tanzania, Peru, Egypt, China, and Saudi Arabia. Several studies suggested that mean sperm density and trends in semen quality may vary considerably, even within small areas (29,30), so that it would have been desirable to stratify studies into narrower geographic categories if sufficient data had been available. Unfortunately, because many of these countries contributed only one study, it was not possible to use narrower geographic strata.

Abstinence time is known to be strongly related to sperm density (31–33). In this analysis, when abstinence time was added to a linear model that included all other variables, the magnitude of the slope decreased by 10.6%, suggesting moderate confounding. Although the inclusion of abstinence time in the model appears to have reduced confounding to some extent, control for this variable was undoubtedly incomplete because less than one-third of these studies included reported abstinence times. An additional 49% of studies noted that abstinence times were restricted by study protocol but, as has been demonstrated, these protocols are only advisory. Auger et al. (2) noted that only 66% of men adhered to the protocol-specified abstinence time of 3–5 days. On the other hand, to account for the observed decline in sperm density, abstinence time would have had to decline appreciably over the study period. The evidence for this is not strong; studies with longer abstinence times (none < 3

Table 5. Comparison of multiple regression models from Swan et al. (13) ($n = 56$) and the current analysis ($n = 101$).[a]

Model	Adjusted R^2	Region	Slope within region	Interaction beta	Interaction p-Value
Linear Swan et al. (13)	0.80	United States	−1.50 (−1.90−−1.10)	Referent	
		Europe/Australia	−3.13 (−4.96−−1.30)	−1.63	0.08
		Other countries	1.56 (−1.00−4.12)	3.06	0.03
Current analysis	0.61	North America	−0.80 (−1.37−−0.24)	Referent	
		Europe	−2.35 (−3.66−−1.05)	−1.55	0.03
		Other countries	−0.21(−2.30−1.88)	0.60	0.56
Spline Swan et al. (13)	0.79	United States < 1970	−1.52 (−2.37−−0.66)	Referent	
		United States ≥ 1970	−1.47 (−3.00−0.06)	0.04	0.97
		Europe/Australia	−3.12 (−4.99−−1.26)	−1.61	0.13
		Other countries	1.56 (−1.03−4.16)	3.08	0.04
Current analysis	0.60	North America <1970	−0.93 (−1.81−−0.05)	Referent	
		North America ≥ 1970	−0.55 (−2.00−0.89)	0.37	0.71
		Europe	−2.32 (−3.64−−1.00)	−1.39	0.09
		Other countries	−0.25 (−2.37−1.86)	0.68	0.52
Step Swan et al. (13)	0.72	United States < 1977	106.7 (91.0−122.5)	—	
		United States ≥ 1970	67.7 (55.9−79.5)	—	
		Europe/Australia	75.0 (60.0−90.0)	—	
		Other countries	58.3 (46.0−70.7)	—	
Current analysis	0.57	North America < 1970	137.9 (116.6−159.3)	—	
		North America ≥ 1970	113.2 (95.9−130.6)	—	
		Europe	120.1 (103.6−136.6)	—	
		Other countries	104.0 (84.7−123.4)	—	

[a]Controlled for proven fertility, abstinence time, age, specimen collection method, study goal and interaction of region and study year [Swan et al. (13)]; and proven fertility, abstinence time, age, specimen collection method, method of counting sperm, whether study was included by Carlsen et al. (1); and interaction of region and study year (current analysis).

days) were published only slightly earlier than those that included some abstinence times < 3 days (1976 vs. 1983).

After controlling for abstinence time and other covariates, the addition of age to the model increased the magnitude of the slope by only 1.2%. However, we found little evidence that age is an important predictor of sperm density. Information was quite incomplete for this variable. Twenty-five studies contained no information on age, and these tended to be older studies (mean publication year 1962). For the remaining studies, many only included an age range, so that we were only able to categorize age into broad categories. Nevertheless, we chose to retain this variable in the model for comparability to other analyses.

The current analysis suggests that the previously reported trends have continued, at least until 1996. We have also shown that the studies initially used by Carlsen et al. (1) did not represent a biased selection of the English language literature. Nevertheless, it is likely that neither this publication nor further statistical analyses of historical data will resolve the continuing debate over declining sperm counts. Critics will continue to challenge the reliability of historical data, and most will agree that residual confounding, which may be appreciable, cannot be completely eliminated.

The entire issue of declining sperm count has gained in importance because of the recognition of several other trends that reflect a decline in male reproductive health.

Testicular cancer incidence has increased significantly for at least the past 20 years in most of the Caucasian populations that have been studied (30,34,35). Trends in rates of cryptorchidism are consistent with those for testicular cancer, for which cryptorchidism is a significant risk factor (30). These increases in rates of testicular cancer and male genital tract abnormalities, like decreasing sperm density, have primarily been seen in Western countries. Several authors have suggested that these trends, together with decreases in semen quality, may reflect a more generalized increase in testicular dysfunction (30,36,37). Although few of these trend studies have examined possible causes, common environmental exposures are plausible. If environmental factors have produced some, or all, of the temporal changes in sperm density, the regional differences that have been reported in semen quality, even within countries (6,38), may also reflect variation in these environmental factors.

Studies that examine differences in semen quality between geographically diverse cohorts may help identify such factors. An ongoing network of international studies, begun in 1997, was designed to address this question. In these collaborative prospective studies, the use of common study protocols, analytic methods, and quality control procedures should minimize extraneous interstudy differences. These studies should provide unbiased estimates of variability among cities that have been reported to differ widely in semen quality, provide baseline levels of male biomarkers for future studies, and generate hypotheses of environmental causes of variation in these parameters.

Appendix 1. Studies not included by Carlsen et al. (1).

Author (reference)	Year	Country	Sample size	Mean sperm density
Aitden et al. (39)	1982	United Kingdom	35	129.30
Assennato et al. (40)	1987	Italy	18	85.00*
Belding (41)	1934	United States	15	119.00
Cooper et al. (42)	1991	Germany	25	94.20
Cottell and Harrison (43)	1995	Ireland	10	159.90
Davidson (44)	1949	United Kingdom	15	143.00
De Castro and Mastrorocco (45)	1984	Brazil	501	66.20
De Castro et al. (46)	1990	Brazil	1,890	37.00
Dougherty et al. (47)	1981	United States	132	83.00
el Shoura et al. (48)	1995	Saudi Arabia	50	116.04
Eskenazi et al. (49)	1991	United States	48	87.00
Fariss et al. (50)	1981	United States	112	75.00
Fedder et al. (51)	1993	Denmark	42	80.00
Figa-Talamanca et al. (52)	1996	Italy	50	62.00
Frick et al. (53)	1981	Finland	13	43.90
Glass et al. (54)	1979	United States	22	61.70
Jelnes (55)	1988	Denmark	68	70.20*
Jensen et al. (56)	1996	Denmark	141	69.20
Kolon et al. (57)	1992	United States	10	132.80
Levine et al. (58)	1992	United States	142	70.74
Milby and Whorton (59)	1980	United States	90	93.00*
Nnatu (60)	1991	Nigeria	21	64.40
Noack-Fuller et al. (61)	1993	Germany	22	86.50
Ratcliffe et al. (62)	1987	United States	43	68.72
Richardson et al. (63)	1984	United Kingdom	47	130.60
Rogers et al. (64)	1979	United States	21	114.00
Rosenberg et al. (65)	1985	United States	71	99.20
Saaranen et al. (66)	1987	Finland	190	141.07
Schrader et al. (67)	1988	United States	45	47.43
Shaarawy and Mahmoud (68)	1982	Egypt	45	77.50
Sheriff (69)	1987	Libya	10	51.40
Sheriff and Legnain (70)	1992	Libya	1,250	85.00
Spira (71)	1984	France	52	92.30
Stankovic and Mikac-Devic (72)	1976	Yugoslavia	67	55.20
Sugkraroek et al. (73)	1991	Canada	20	128.50
Venable et al. (74)	1980	United States	63	113.53
Vignon et al. (75)	1991	France	20	76.50
Vogt et al. (76)	1986	Germany	239	62.69
Wallace et al. (77)	1993	United Kingdom	28	86.50
Wang and Yeung (78)	1980	Hong Kong	15	88.90
Ward et al. (79)	1984	United States	11	87.40
Weidner et al. (80)	1991	Germany	42	52.00*
Weyandt et al. (81)	1996	United States	31	38.00
Wickings et al. (82)	1983	Germany	25	84.70
Wyrobek et al. (83)	1981	United States	26	66.20
Wyrobek et al. (84)	1981	United States	34	128.70
Zhong et al. (85)	1990	China	19	96.60

*Arithmetic mean estimated from median or geometric mean (see text).

REFERENCES AND NOTES

1. Carlsen E, Giwercman A, Keiding N, Skakkebaek N. Evidence for decreasing quality of semen during past 50 years. Br Med J 305:609–613 (1992).
2. Auger J, Kunstmann JM, Czyglik F, Jouannet P. Decline in semen quality among fertile men in Paris during the past 20 years. N Engl J Med 332:281–285 (1995).
3. Irvine S, Cawood E, Richardson D, MacDonald E, Aitken J. Evidence of deteriorating semen quality in the United Kingdom: birth cohort study in 577 men in Scotland over 11 years. Br Med J 312:467–471 (1996).
4. Van Waeleghem K, De Clercq N, Vermeulen L, Schoonjans F, Comhaire F. Deterioration of sperm quality in young healthy Belgian men. Hum Reprod 11(2):325–329 (1996).
5. Vierula M, Niemi M, Keiski A, Saaranen M, Saarikoski S, Suominen J. High and unchanged sperm counts of Finnish men. Int J Androl 19(1):11–17 (1996).
6. Fisch H, Goluboff ET, Olson JH, Feldshuh J, Broder SJ, Barad DH. Semen analyses in 1,283 men from the United States over a 25-year period: no decline in quality. Fertil Steril 65(5):1009–1014 (1996).
7. 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 65(5):1015–1020 (1996).
8. Olsen GW, Bodner KM, Ramlow JM, Ross CE, Lipshultz LI. Have sperm counts been reduced 50 percent in 50 years? A statistical model revisited. Fertil Steril 63(4):887–893 (1995).
9. Lerchl A, Nieschlag E. Decreasing sperm counts? A critical (re)view. Exp Clin Endocrinol Diabetes 104(4):301–307 (1996).

10. Brake A, Krause W. Decreasing quality of semen [Letter]. Br Med J 305:1498 (1992).

11. Farrow S. Falling sperm quality: fact or fiction? Br Med J 309:1–2 (1994).

12. Bromwich P, Cohen J, Stewart I, Walker A. Decline in sperm counts: an artefact of changed reference range of "normal"? Br Med J 309(6946):19–22 (1994).

13. Swan SH, Elkin EP, Fenster L. Have sperm densities declined? A reanalysis of global trend data. Environ Health Perspect 105:1228–1232 (1997).

14. Swan SH, Elkin EP. Declining semen quality: can the past inform the present? Bioessays 21(7):614–621 (1999).

15. Varnek J. Spermaens maegde. In: Spermaundersogeser ved sterilitet:Med specielt henblik pa spermiernes morfologi Aarhus, Denmark:Universitetsforlaget, 1944;42–52.

16. Robles GG. Estudio del liquido espermatico. Arch Peruanos Patol Clin 1:615–661 (1947).

17. Sturde H-C, Glowania HJ, Bohm K. Vergleichende ejaculatuntersuchungen bei mannern aus sterilen und fertilen ehen. Arch Dermatol Forsch 241:426–437 (1971).

18. Santomauro AG, Sciarra JJ, Varma AO. A clinical investigation of the role of the semen analysis and postcoital test in the evaluation of male infertility. Fertil Steril 23:245–251 (1972).

19. Bahamondes L, Abdelmassih R, Dachs JN. Survey of 185 sperm analyses of fertile men in an infertility service. Int J Androl 2:526–533 (1979).

20. Polakoski KL, Zahler WL, Paulson JD. Demonstration of proacrosin and quantitation of acrosin in ejaculated human spermatozoa. Fertil Steril 28(6):668–673 (1977).

21. Lewis EL, Brazil CK, Overstreet JW. Human sperm function in the ejaculate following vasectomy. Fertil Steril 42(6):895–898 (1984).

22. SAS Institute Inc. SAS/STAT User's Guide, Version 6. 4th ed. Cary, NC:SAS Institute Inc., 1989.

23. Rothman KJ, Greenland S. Introduction to stratified analysis. In: Modern Epidemiology. Philadelphia, PA:Lippincott-Raven, 1998:253–279.

24. Zavos PM, Goodpasture JC. Clinical improvements of specific seminal deficiencies via intercourse with a seminal collection device versus masturbation. Fertil Steril 51(1):190–193 (1989).

25. Levine RJ, Bordson BL, Mathew RM, Brown MH, Stanley JM, Star TB. Deterioration of semen quality during summer in New Orleans. Fertil Steril 49(5):900–907 (1988).

26. Laboratory Manual for the Examination of Human Semen and Semen-Cervical Mucus Interaction: Based on Consultations Held within the WHO Special Programme of Research, Development and Research Training in Human Reproduction. Singapore:Press Concern, 1980.

27. WHO Laboratory Manual for the Examination of Human Semen and Semen-Cervical Mucus Interactions. 2nd ed. Cambridge, UK:Cambridge University Press, 1987.

28. Makler A. The improved ten-micrometer chamber for rapid sperm count and motility evaluation. Fertil Steril 33(3):337–338 (1980).

29. Fisch H, Goluboff ET. Geographic variations in sperm counts: a potential cause of bias in studies of semen quality. Fertil Steril 65(5):1044–1046 (1996).

30. Toppari J, Larsen JC, Christiansen P, Giwercman A, Grandjean P, Guillette LJ Jr, Jégou B, Jensen TK, Jouannet P, Keiding N, et al. Male reproductive health and environmental xenoestrogens. Environ Health Perspect 104(suppl 4):741–803 (1996).

31. Macleod J, Gold RZ. The male factor in fertility and infertility: spermatozoan counts in 1000 men of known fertility and in 1000 cases of infertile marriage. J Urol 66:436–449 (1951).

32. Schwartz D, Laplanche A, Jouannet P, David G. Withinsubject variability of human semen in regard to sperm count, volume, total number of spermatozoa and length of abstinence. J Reprod Fertil 57:391–395 (1979).

33. Magnus O, Tollefsrud A, Abyholm T, Purvis K. Effects of varying the abstinence period in the same individuals on sperm quality. Arch Androl 26:199–203 (1991).

34. Bergstrom R, Adami HO, Mohner M, Zatonski W, Storm H, Ekbom A, Tretli S, Teppo L, Akre O, Hakulinen T. Increase in testicular cancer incidence in six European countries: a birth cohort phenomenon. J Natl Cancer Inst 88(11):727–733 (1996).

35. Adami H, Bergstrom R, Mohner M, Zatonski W, Storm H, Ekbom H, Tretli S, Teppo L, Ziegler H, Rahu M. Testicular cancer in nine northern European countries. Int J Cancer 59:33–38 (1994).

36. Sharpe RM, Skakkebaek NE. Are oestrogens involved in falling sperm counts and disorders of the male reproductive tract? Lancet 341(8857):1392–1395 (1993).

37. Giwercman A, Carlsen E, Keiding N, Skakkebaek NE. Evidence for increasing incidence of abnormalities of the human testis: a review. Environ Health Perspect 101(suppl 2):65–71 (1993).

38. Auger J, Jouannet P. Evidence for regional differences of semen quality among fertile French men. Federation Francaise des Centres d'Etude et de Conservation des Oeufs et du Sperme humains. Hum Reprod 12(4):740–745 (1997).

39. Aitken RJ, Best FS, Richardson DW, Djahanbakhch O, Lees MM. The correlates of fertilizing capacity in normal fertile men. Fertil Steril 38(1):68–76 (1982).

40. Assennato G, Paci C, Baser ME, Molinini R, Candela RG, Altamura BM, Giorgino R. Sperm count suppression without endocrine dysfunction in lead-exposed men. Arch Environ Health 42(2):124–127 (1987).

41. Belding DL. Fertility in the male; technique of spermatozoa count. Am J Obstet Gynecol 27:25–31 (1934).

42. Cooper TG, Jockenhovel F, Nieschlag E. Variations in semen parameters from fathers. Hum Reprod 6:859–866 (1991).

43. Cottell E, Harrison RF. The value of subcellular elemental analysis in the assessment of human spermatozoa. Hum Reprod 10(12):3186–3189 (1995).

44. Davidson HA. Male subfertility: interim report of 3,182 cases. Br Med J 2:1328–1332 (1949).

45. de Castro MP, Mastrorocco DA. Reproductive history and semen analysis in prevasectomy fertile men with and without varicocele. J Androl 5(1):17–20 (1984).

46. de Castro M, Jeyendran RS, Zaneveld LJ. Hypo-osmotic swelling test: analysis of prevasectomy ejaculates. Arch Androl 24(1):11–16 (1990).

47. Dougherty RC, Whitaker MJ, Tang SY, Bottcher R, Keller M, Kuehl DW. Sperm density and toxic substances: a potential key to environmental health hazards. In: Environmental Health Chemistry: The Chemistry of Environmental Agents of Potential Human Hazards (McKinney JD, ed). Ann Arbor, MI:Ann Arbor Science Publishers, 1981:263–278.

48. el Shoura SM, Abdel AM, Ali ME, el Said MM, Ali KZ, Kemeir MA, Raoof AM, Allam M, Elmalik EM. Deleterious effects of khat addiction on semen parameters and sperm ultrastructure. Hum Reprod 10(9):2295–2300 (1995).

49. Eskenazi B, Wyrobek AJ, Fenster L, Katz DF, Sadler M, Lee J, Hudes M, Rempel DM. A study of the effect of perchloroethylene exposure on semen quality in dry cleaning workers. Am J Ind Med 20(5):575–591 (1991).

50. Fariss BL, Fenner DK, Plymate SR, Brannen GE, Jacob WH, Thomason AM. Seminal characteristics in the presence of a varicocele as compared with those of expectant fathers and prevasectomy men. Fertil Steril 35(3):325–327 (1981).

51. Fedder J, Askjaer SA, Hjort T. Nonspermatozoal cells in semen: relationship to other semen parameters and fertility status of the couple. Arch Androl 31(2):95–103 (1993).

52. Figa-Talamanca I, Cini C, Varricchio GC, Dondero F, Gandini L, Lenzi A, Lombardo F, Angelucci L, Di Grezia R, Patacchioli FR. Effects of prolonged autovehicle driving on male reproduction function: a study among taxi drivers. Am J Ind Med 30(6):750–758 (1996).

53. Frick J, Danner C, Joos H, Kunit G, Luukkainen T. Spermatogenesis in men treated with subcutaneous application of levonorgestrel and estrone rods. J Androl 2(6):331–338 (1981).

54. Glass RI, Lyness RN, Mengle DC, Powell KE, Kahn E. Sperm count depression in pesticide applicators exposed to dibromochloropropane. Am J Epidemiol 109(3):346–351 (1979).

55. Jelnes JE. Semen quality in workers producing reinforced plastic. Reprod Toxicol 2(3-4):209–212 (1988).

56. Jensen TK, Giwercman A, Carlsen E, Scheike T, Skakkebaek NE. Semen quality among members of organic food associations in Zealand, Denmark [Letter]. Lancet 347(9018):1844 (1996).

57. Kolon TF, Philips KA, Buch JP. Custom cryopreservation of human semen. Fertil Steril 58(5):1020–1023 (1992).

58. Levine RJ, Brown MH, Bell M, Shue F, Greenberg GN, Bordson BL. Air-conditioned environments do not prevent deterioration of human semen quality during the summer. Fertil Steril 57(5):1075–1083 (1992).

59. Milby TH, Whorton D. Epidemiological assessment of occupationally related, chemically induced sperm count suppression. J Occup Med 22(2):77–82 (1980).

60. Nnatu SN, Giwa-Osagie OF, Essien EE. Effect of repeated semen ejaculation on sperm quality. Clin Exp Obstet Gynecol 18(1):39–42 (1991).

61. Noack-Fuller G, De Beer C, Seibert H. Cadmium, lead, selenium, and zinc in semen of occupationally unexposed men. Andrologia 25(1):7–12 (1993).

62. Ratcliffe JM, Schrader SM, Steenland K, Clapp DE, Turner T, Hornung RW. Semen quality in papaya workers with long term exposure to ethylene dibromide. Br J Ind Med 44(5):317–326 (1987).

63. Richardson DW, Aitken RJ, Loudon NB. The functional competence of human spermatozoa recovered after vasectomy. J Reprod Fertil 70(2):575–579 (1984).

64. Rogers BJ, Van Campen H, Ueno M, Lambert H, Bronson R, Hale R. Analysis of human spermatozoal fertilizing ability using zona-free ova. Fertil Steril 32(6):664–670 (1979).

65. Rosenberg MJ, Wyrobek AJ, Ratcliffe J, Gordon LA, Watchmaker G, Fox SH, Moore DH, Hornung RW. Sperm as an indicator of reproductive risk among petroleum refinery workers. Br J Ind Med 42(2):123–127 (1985).

66. Saaranen M, Suonio S, Kauhanen O, Saarikoski S. Cigarette smoking and semen quality in men of reproductive age. Andrologia 19(6):670–676 (1987).

67. Schrader SM, Turner TW, Breitenstein MJ, Simon SD. Longitudinal study of semen quality of unexposed workers. I. Study overview. Reprod Toxicol 2(3-4):183–190 (1988).

68. Shaarawy M, Mahmoud KZ. Endocrine profile and semen characteristics in male smokers. Fertil Steril 38(2):255–257 (1982).

69. Sheriff DS. Semen analyses in Hansen's disease. Trans R Soc Trop Med Hyg 81(1):113–114 (1987).

70. Sheriff DS, Legnain M. Evaluation of semen quality in a local Libyan population. Indian J Physiol Pharmacol 36(2):83–87 (1992).

71. Spira A. Seasonal variations of sperm characteristics. Arch Androl 12(suppl):23–28 (1984).

72. Stankovic H, Mikac-Devic D. Zinc and copper in human semen. Clinica Chimica Acta 70(1):123–126 (1976).

73. Sugkraroek P, Kates M, Leader A, Tanphaichitr N. Levels of cholesterol and phospholipids in freshly ejaculated sperm and Percoll-gradient-pelletted sperm from fertile and unexplained infertile men. Fertil Steril 55(4):820–827 (1991).

74. Venable JR, McClimans CD, Flake RE, Dimick DB. A fertility study of male employees engaged in the manufacture of glycerine. J Occup Med 22(2):87–91 (1980).

75. Vignon F, Le Faou A, Montagnon D, Pradignac A, Cranz C, Winiszewsky P, Pinget M. Comparative study of semen in diabetic and healthy men. Diabete Metab 17(3):350–354 (1991).

76. Vogt HJ, Heller WD, Borelli S. Sperm quality of healthy smokers, ex-smokers, and never-smokers. Fertil Steril 45(1):106–110 (1986).

77. Wallace EM, Gow SM, Wu FC. Comparison between testosterone enanthate-induced azoospermia and oligozoospermia in a male contraceptive study. I: Plasma luteinizing hormone, follicle stimulating hormone, testosterone, estradiol, and inhibin concentrations. J Clin Endocrinol Metab 77(1):290–293 (1993).

78. Wang C, Yeung KK. Use of low-dosage oral cyproterone acetate as a male contraceptive. Contraception 21(3):245–272 (1980).

79. Ward JB Jr, Hokanson JA, Smith ER, Chang LW, Pereira MA, Whorton EB Jr, Legator MS. Sperm count, morphology and fluorescent body frequency in autopsy service workers exposed to formaldehyde. Mutat Res 130(6):417–424 (1984).

80. Weidner W, Jantos C, Schiefer HG, Haidl G, Friedrich HJ. Semen parameters in men with and without proven chronic prostatitis. Arch Androl 26(3):173–183 (1991).

81. Weyandt TB, Schrader SM, Turner TW, Simon SD. Semen analysis of military personnel associated with military duty assignments. Reprod Toxicol 10(6):521–528 (1996).

82. Wickings EJ, Freischem CW, Langer K, Nieschlag E. Heterologous ovum penetration test and seminal parameters in fertile and infertile men. J Androl 4(4):261–271 (1983).

83. Wyrobek AJ, Watchmaker G, Gordon L, Wong K, Moore D II, Whorton D. Sperm shape abnormalities in carbaryl-exposed employees. Environ Health Perspect 40:255–265 (1981).

84. Wyrobek AJ, Brodsky J, Gordon L, Moore DH, Watchmaker G, Cohen EN. Sperm studies in anesthesiologists. Anesthesiology 55(5):527–532 (1981).

85. Zhong CQ, Lui QL, Tang YJ, Wang Y, Shi FJ, Qian SZ. Study on sperm function in men long after cessation of gossypol treatment. Contraception 41(6):617–622 (1990).