Geographic Differences in Semen Quality of Fertile U.S. Males

Shanna H. Swan,[1] Charlene Brazil,[2] Erma Z. Drobnis,[3] Fan Liu,[1] Robin L. Kruse,[1] Maureen Hatch,[4] J. Bruce Redmon,[5] Christina Wang,[6] James W. Overstreet,[2] and The Study for Future Families Research Group

[1]Department of Family and Community Medicine, University of Missouri-Columbia School of Medicine, Columbia, Missouri, USA; [2]University of California, Davis, California, USA; [3]Department of Obstetrics and Gynecology, University of Missouri-Columbia, Columbia, Missouri, USA; [4]Mount Sinai School of Medicine, New York, New York, USA; [5]Departments of Medicine and Urologic Surgery, University of Minnesota, Minneapolis, Minnesota, USA; [6]Harbor-UCLA Medical Center and Research and Education Institute, Torrance, California, USA

Although geographic variation in semen quality has been reported, this is the first study in the United States to compare semen quality among study centers using standardized methods and strict quality control. We evaluated semen specimens from partners of 512 pregnant women recruited through prenatal clinics in four U.S. cities during 1999–2001; 91% of men provided two specimens. Sperm concentration, semen volume, and motility were determined at the centers, and morphology was assessed at a central laboratory. Study protocols were identical across centers, and quality control was rigorously maintained. Sperm concentration was significantly lower in Columbia, Missouri, than in New York, New York; Minneapolis, Minnesota; and Los Angeles, California. Mean counts were 58.7, 102.9, 98.6, and 80.8 × 10^6/mL (medians 53.5, 88.5, 81.8, and 64.8 × 10^6/mL) in Missouri, New York, Minnesota, and California, respectively. The total number of motile sperm was also lower in Missouri than in other centers: 113, 196, 201, and 162 × 10^6 in Missouri, New York, Minnesota, and California, respectively. Semen volume and the percent morphologically normal sperm did not differ appreciably among centers. These between-center differences remained significant in multivariate models that controlled for abstinence time, semen analysis time, age, race, smoking, history of sexually transmitted disease, and recent fever (all p-values < 0.01). Confounding factors and differences in study methods are unlikely to account for the lower semen quality seen in this mid-Missouri population. These data suggest that sperm concentration and motility may be reduced in semirural and agricultural areas relative to more urban and less agriculturally exposed areas. Key words: agriculture, geography, semen quality, sperm concentration, sperm morphology, sperm motility. Environ Health Perspect 111:414–420 (2003). doi:10.1289/ehp.5927 available via http://dx.doi.org/ [Online 11 November 2002]

Historically, semen parameter studies have included highly selected and nonrepresentative subgroups such as compensated sperm donors, prevasectomy patients, or infertility clinic populations. Moreover, measures of semen quality are very sensitive to the methods of semen collection (including abstinence time) and analysis, which vary significantly among study sites. Further, most analyses of temporal trends and geographic variation in semen parameters have been retrospective and subject to confounding by factors such as smoking or recent high fever that cannot be well controlled retrospectively. These studies have been conducted almost exclusively at andrology centers, which are usually located in urban areas, primarily in Western Europe and North America.

Nonetheless, over the past decade several authors have reported large geographic differences between cities in mean sperm concentration. For example, an international study of testosterone-induced azoospermia found that mean pretreatment sperm concentrations of normal men in nine countries ranged from 52.1 × 10^6/mL in Bangkok, Thailand, to 103.5 × 10^6/mL in Melbourne, Australia [World Health Organization (WHO) Task Force on Methods of Regulation of Male Fertility 1996]. A wide range of sperm concentration was also reported in eight cities in France (Auger and Jouannet 1997). Several recent studies suggest that wide variation is also present among cities in the United States. Wittmaack and Shapiro (1992) examined sperm concentration between 1978 and 1987 in Madison, Wisconsin; mean sperm concentration during this time was approximately 80 × 10^6/mL. Paulsen et al. (1996) reported a geometric mean of about 50 × 10^6/mL in Seattle, Washington, during 1972–1993. A recent study in California (Fenster et al. 1997) found a median sperm concentration of 64 × 10^6/mL. Fisch and Goluboff (1996) reported large differences in mean sperm concentration in prevasectomy patients from Los Angeles, California; Minneapolis, Minnesota; and New York City, New York, with low concentration in Los Angeles compared with Minneapolis and New York City (72.7 vs. 100.8 and 131.5 × 10^6/mL, respectively). Because these retrospective studies used data collected under a variety of protocols, differences in population selection or methods of semen analysis may have contributed to these differences.

Recent multicenter studies have sought to eliminate many limitations of earlier studies by standardizing methods and populations. Recognizing that carefully controlled, prospective studies of semen parameters are needed, several multicenter national and international studies have been underway since 1997. The International Study of Semen Quality in Partners of Pregnant Women was recently completed in Europe (Jorgenson et al. 2001). This study found significant differences in mean sperm count and other semen parameters between fertile men recruited in Copenhagen, Denmark; Paris, France; Edinburgh, Scotland: and Turku, Finland. For example, sperm concentration in Copenhagen was only 74% that in Turku. The observed differences were not changed appreciably by adjustment for age, abstinence time, and season.

The ongoing Study for Future Families was designed in collaboration with the European study, so that meaningful comparisons can be made between U.S. and European

centers. The Study for Future Families examines semen quality and other reproductive parameters of fertile couples recruited at prenatal clinics in four cities in the north, east, west and south-central United States, using methods for clinical examination, data collection, and semen analysis that are identical across U.S. centers and consistent with those used in the European study.

Materials and Methods

Study subjects. In the Study for Future Families, a 4-year study funded by the National Institute of Environmental Health Sciences, we have been recruiting women at prenatal clinics affiliated with university hospitals in Los Angeles (Harbor-UCLA and Cedars-Sinai); Minneapolis (University of Minnesota Health Center); Columbia, Missouri (University Physicians); and New York (Mt. Sinai School of Medicine) since September 1999. We use a standardized recruitment protocol at each center to minimize between-center differences. Any woman who keeps her prenatal appointment at a study clinic during a recruitment session is a potential subject, and the outcome of every potential subject (eligibility and level of participation, if eligible) is determined and recorded in the potential subject database. The couple is eligible unless the woman or her partner is < 18 years of age; the pregnancy was medically assisted; either partner does not read and speak Spanish or English; the father is unavailable or unknown; the couple does not plan to stay in the area (because couples planning to move out of area would be unlikely to complete their study participation); the pregnancy is medically threatened; or either partner is incompetent or a prisoner. We ask eligible women to take home study information and a recruitment video to review with their partners. If the couple agrees to participate, the man completes a questionnaire, receives a physical examination, and gives a blood sample, a urine sample and two semen samples. The woman completes a questionnaire and gives blood and urine samples. All study instruments (including questionnaires, mini-questionnaires, letters, and instructions for the man) were translated into Spanish and back-translated for accuracy. The instructional video was also produced in both English and Spanish. Subjects are offered monetary compensation, the amount reflecting the cost of living in the study area. In this communication we report semen analysis results from the 512 men who completed participation by 15 November 2001.

For the present study, recruiters asked eligible couples who refused to participate to answer a very brief mini-questionnaire that includes demographics, history of infertility, and time to pregnancy (TTP). They also asked a sample of study participants to answer the same questions. We compared responses between refusals and study subjects to examine selection bias. This issue was also examined by comparing questionnaires of subjects who gave a semen sample with those of men who agreed to participate in the study but preferred not to give a semen sample.

The number of subjects varied by center and was particularly low in New York, where the closure of the Mt. Sinai andrology center in the second study year resulted in a shortened period of recruitment. However, because New York results on study-wide quality control samples were in close agreement with other centers, and coefficients of variation for New York technicians were low, New York data could be meaningfully compared with other centers, despite small numbers. We also conducted an analysis that examined the impact of excluding New York subjects on the estimates of differences in semen quality among the remaining centers.

Semen collection and analysis. We requested that subjects observe a 2- to 5-day abstinence period before providing a semen sample. Prior to each of the two visits, which were approximately 3 weeks apart, we mailed instructions regarding specimen collection, including a schedule to assist the subject in timing his last ejaculation prior to the visit. At the time of the visit we stressed the importance of accurately reporting the actual abstinence period and assured men that their sample would not be rejected if they deviated from the recommended protocol. At the study visit men collected semen samples by masturbation at the clinic, and these were analyzed within 45 min of collection.

We determined sperm concentration for each of the two samples using a μ-Cell disposable counting chamber (Conception Technologies, San Diego, CA) and, for the first sample only, a hemacytometer (Improved Neubauer; Hauser Scientific Inc., Horsham, PA). Regardless of the counting method, sperm concentration was estimated for each sample as the mean of two readings, unless these differed by > 10%, in which case a third reading was taken and it was estimated by the median of the three counts. Ejaculate volumes were estimated by specimen weight, assuming a semen density of 1.0 g/mL. For this calculation each container was preweighed and the weight (written on the container) was subtracted from the weight of the container plus sample.

In the present analysis the percent motile sperm was counted in a μ-Cell chamber (Overstreet and Brazil 1997) and refers to the percentage of sperm with any flagellar movement, whether twitching or progressive. We calculated the total motile count (TMC) by multiplying the sperm concentration by the semen volume; values obtained by each of the two sperm-counting methods were used to calculate TMC. Determinations of motility by the methods recommended by the WHO (1999) were made on the first sample only; these data are not discussed here.

Seminal smears were prepared at the clinical centers and shipped to the Andrology Coordinating Center (ACC) at the University of California-Davis for Papanicalou staining, analysis, and storage. Sperm morphology was assessed by a single technician using the strict morphology method (WHO 1999) and by a second technician using more traditional 1987 WHO criteria (WHO 1987). For each determination, 100 consecutive sperm were scored in each of two randomly selected areas of the slide and the percentage with normal morphology was determined. Under strict criteria for assessing morphology [the only method reported here and recommended by the WHO (Guzick et al. 2001; WHO 1999)], only sperm with absolutely no defects were classified as normal.

In addition to the primary measures of semen quality (sperm concentration, volume, percent morphologically normal sperm, and percent motile sperm), we analyzed two derived semen parameters: total count (TC; sperm concentration × volume), and total motile count (TMC; TC × percent motile). TC and TMC were calculated using both μ-Cell and hemacytometer estimates of concentration.

Technicians from each study site attended a week-long training session at the ACC and had to be certified by passing a proficiency test before conducting any semen analyses for this study. The ACC also conducted quarterly quality control testing. Test results were reported to the ACC, where within- and between-technician variability were assessed. A coefficient of variation (CV) was calculated for each technician based on the average of four blind readings of each ejaculate, and these were averaged to obtain the intratechnician CV for each technique. Throughout the course of this study, all andrology technicians achieved CVs of ≤ 15%. The technicians' average values were within 15% of standard values for all semen parameters throughout the course of the study, except for hemacytometer counts, which were within 17% of standard values.

Statistical analysis. The primary outcomes of interest in these analyses were between-center differences in semen parameters, which we estimated in two ways. First, we calculated simple (untransformed and unadjusted) means, as these are easy to interpret and to compare with published studies. We report unadjusted sperm counts based on one sample per man, obtained by hemacytometer, the most frequently used method for counting sperm (Brouwer et al. 1998). We

also include counts obtained using the μ-Cell chamber for comparison.

Because sperm concentration, semen volume, TC, and TMC follow markedly skewed (non-normal) distributions, they must be transformed before analysis. We transformed the data using logarithm (base 10), which is generally recommended (Berman et al. 1996) for transformation of skewed semen parameters. We then used multivariate models to adjust for covariates of semen quality that appeared to confound these between-center comparisons. Finally, we back-transformed the regression coefficients for logarithmically transformed variables for ease of interpretation.

Because most men (85%) provided two specimens and because of the expected correlation between semen samples, mixed models that account for repeated measures were fit (Laird and Ware 1982; SAS Institute 2001; Zeger and Liang 1986) assuming a compound symmetry covariance structure (equivalent to assuming that all samples within a man are equally correlated). We used these models to analyze all semen characteristics determined on both the first and second samples. Concentrations by hemacytometer (and TC and TMC based on hemacytometer), which were only available for the man's first sample, were analyzed using a general linear model (SAS Institute 2001). We then compared the between-center differences in semen quality based on simple (unadjusted) means to the adjusted estimates obtained from these multivariate models.

We compared a number of self-reported variables across study centers and examined their relationships to semen characteristics. These include age, race, smoking, education, body mass index, fever in the 3 months before study entry, use of steroids, history of infertility, history of sexually transmitted disease (STD), cryptorchidism, and other genital problems. Characteristics of the semen sample and analysis that were examined include abstinence time, season (January–March, April–June, July–September, and October–December), time from sample collection to start of semen evaluation, and time to perform the semen evaluation. We excluded samples with missing or unknown abstinence times, or with reported abstinence of < 2 hr or > 10 days. Selection of covariates for the final model was based on their importance in the literature, biological plausibility, sufficient numbers within strata, and evidence of some effect on between-center comparisons.

Results

At the time the data set was created for this analysis, we had identified 4,825 potential subjects, of whom 33% were ineligible. Primary reasons for ineligibility include more than 36 weeks pregnant (38%), partner not available

(18%), conception medically assisted (10%), not literate in English or Spanish (8%), either partner under 18 years of age (7%), or not pregnant (7%). Among eligible subjects, 55% refused participation (49, 63, 46, and 60% in Missouri, California, Minnesota, and New York, respectively), and 12% of subjects were lost to follow-up. Among eligible subjects who refused participation or were lost to follow-up, 40% completed a mini-questionnaire (45, 29, 53, and 34% in Missouri, California, Minnesota, and New York, respectively). At the time of this analysis, an additional 11% of subjects had expressed interest in the study or had begun but not yet completed participation. We compared questionnaire responses of subjects who would only participate if they were not required to provide a semen sample (19% of completed subjects) with the 512 men who provided one (48) or two (464) semen samples and had completed participation by 15 November 2001. From these 512 men we excluded 19 because of missing or out-of-range abstinence times. As discussed below, these exclusions did not affect study conclusions.

Univariate analyses. After exclusions, 493 men were available for analysis, of whom 410 provided two semen samples an average of 24 days apart. The abstinence time-adjusted mean sperm concentration for these two samples did not differ (*p* = 0.36), and results of both semen evaluations are included in these analyses.

Several population characteristics varied considerably; study populations and sample characteristics at the four centers are summarized in Table 1. Race varied by center; in California only 23% of subjects were white (non-Hispanic) compared with 86% in Minnesota and Missouri. Subjects in California were also less educated (25%

graduated college or technical school compared with 75% in Minnesota and 74% in New York). Age differed among centers, though less markedly; subjects were slightly younger in California (mean 30 years) and somewhat older in New York (mean 36 years). The proportion of men who smoked at least 10 cigarettes per day also varied somewhat by center and ranged from 3% in New York to 13% in Missouri. History of an STD (gonorrhea, chlamydia, or genital warts) was reported by 13% of men, and 3.6% reported a fever (≥ 101°F) in the 3 months before semen collection.

Mean abstinence time was within 6 hr of the study average (78 hr) at all centers. Time from specimen collection to start of semen analysis was also similar across centers and averaged 30 min. The time to conduct the semen evaluation varied somewhat more across centers. This time was shorter for the second semen evaluation (average 62 and 41 min for first and second sample, respectively) because the second evaluation did not include determination of concentration by hemacytometer or evaluation of motility using WHO methods (WHO 1999).

As shown in Table 2, which contains unadjusted semen parameters from each center, mean sperm concentration in Missouri was lower than at all other centers. Mean (hemacytometer) concentration was 38% higher in California than in Missouri. Greater differences were seen comparing Missouri with New York and Minnesota, which were 75 and 67% higher than Missouri, respectively. In this unadjusted comparison, the percent motile sperm was 8–17% higher in other centers relative to Missouri. Mean TMC was higher in all centers, but particularly in New York and Minnesota; compared with Missouri, New York and Minnesota

Table 1. Characteristics of the study population and semen samples by center.[a]

Characteristics	Missouri	California	Minnesota	New York	Total
Population characteristics					
No. of participants	176	124	155	38	493
Mean age (years)	30.7	29.8	32.2	36.1	31.3
Education					
Less than college	42.5	75.2	25.2	26.3	43.8
College/technical school	57.5	24.8	74.8	73.7	56.2
Non-white race	14.2	77.4	14.2	31.6	31.4
Smoking status					
Nonsmoking	79.5	70.5	85.8	81.6	79.4
< 10/day	7.4	26.2	10.3	15.8	13.7
> 10/day	13.1	3.3	3.9	2.6	6.9
Recent fever	4.0	2.4	4.5	2.6	3.6
Steroid use	2.3	4.1	1.9	0.0	2.4
History of STD	11.4	12.9	13.6	15.8	12.8
History of genital disease other than STD	10.2	8.9	9.0	5.3	9.1
History of infertility	0.6	0.0	1.3	2.6	0.8
Sample characteristics[b]					
Mean ejaculation abstinence time (hr)	78	84	72	84	78
Mean time to start of semen analysis (min)	26	28	30	34	28
Mean time to conduct semen analysis (min)	45	52	61	47	52

[a]Reported figures are percents unless otherwise indicated. [b]Average time for the first and second samples.

were 74 and 77% higher, respectively. Semen volume and the percent morphologically normal sperm differed little among centers. Between-center differences in TC were similar to those seen for sperm concentration and are not presented here. In Table 2 we include sperm characteristics determined both by the µ-Cell chamber and hemacytometer to allow for comparisons between estimates obtained by these two methods.

Table 2 also contains the crude (unadjusted) relationships between covariates and semen parameters. However, these relationships may be somewhat misleading, as they are unadjusted for confounding, which can be appreciable. For example, based on these unadjusted estimates, it would appear that semen volume increases with age. In fact, after adjustment for abstinence time and other covariates, semen volume is seen to decrease at older ages (Table 3).

Multivariate models. Of the subject characteristics examined, race, age, smoking, recent fever, and history of STD were retained in final models, as were abstinence time, time from specimen collection to start of semen analysis, and time to conduct the semen evaluation. Genital infections other than STDs, education, and body mass index did not confound these analyses and were not retained in final models. Because steroid use was reported by only 12 men, and a history of infertility by four, these variables could not be examined further. Figure 1 contains (back-transformed) adjusted estimates of center-specific estimates of semen quality. The differences between Missouri and other centers based on these adjusted data are similar to those based on unadjusted means, as can be seen by comparing results from Table 4 with those from Table 1. For example, percent motile sperm was 17% higher in New York than Missouri using unadjusted data, compared with 21% after adjustment. Differences (both adjusted and unadjusted) between Missouri and others centers were somewhat greater when based on hemacytometer counts than on µ-Cell counts. The (adjusted) sperm concentration in Minnesota, for example, was 62% higher than that in Missouri when based on hemacytometer, compared with 45% higher when based on µ-Cell concentration. Thus, the (crude) unadjusted estimates based on µ-Cell concentrations provide somewhat conservative estimates of between-center differences.

Table 3 shows regression coefficients for all covariates in relation to semen parameters. Age was not related to concentration, morphology, or motility, but a strong nonlinear (quadratic) relationship was seen between volume and age

Table 2. Mean semen characteristics (unadjusted, untransformed) by center and covariates.

Center/covariate	Sperm concentration (10^6/mL)		Volume (g)	Percent motile sperm	TMC (10^6)		Percent normal sperm
	Hemacytometer	µ-Cell			Hemacytometer	µ-Cell	
No. of samples	472	901	901	903	471	899	887
Center							
Missouri	58.7	53.4	3.9	48.2	113.0	101.0	10.8
California	80.8	69.0	3.6	54.5	162.2	137.5	12.2
Minnesota	98.6	74.6	3.9	52.1	200.9	152.9	11.4
New York	102.9	75.5	3.3	56.4	196.4	149.7	10.9
Covariate							
Age (years)							
< 25	70.1	60.4	3.4	52.8	127.5	109.1	11.5
25–34	81.7	67.0	3.8	52.4	170.6	138.8	11.4
≥ 35	82.0	65.8	3.8	49.3	151.9	123.8	11.2
Education							
< College	72.5	61.0	3.5	51.8	138.9	114.3	11.5
College/technical school	86.2	69.4	4.0	51.4	175.4	141.8	11.2
White	79.5	65.4	3.9	51.1	160.5	130.8	11.1
Non-white	81.2	66.2	3.5	52.9	156.2	128.6	11.9
Smoking							
Nonsmoking	82.1	67.9	3.9	51.3	165.4	136.2	11.2
< 10 cigarettes/day	77.2	60.3	3.5	53.9	147.8	116.2	12.2
> 10 cigarettes/day	59.7	49.3	3.2	50.8	111.8	87.2	11.0
Recent fever							
Yes	68.4	51.5	3.6	43.3	120.4	95.8	9.8
No	80.4	66.2	3.8	52.0	160.6	131.5	11.4
STDs							
Yes	73.0	57.7	3.7	50.0	140.5	111.2	9.8
No	81.1	66.8	3.8	51.9	162.0	132.9	11.6

Table 3. Summary of adjusted semen characteristics by covariates.

Covariate	Sperm concentration (10^6/mL)		Volume (g)	Percent motile	TMC (10^6)		Percent normal sperm
	Hemacytometer	µ-Cell			Hemacytometer	µ-Cell	
Age[a]	0.029 (0.18)	0.018 (0.32)	0.040 (< 0.001)	0.31 (0.58)	0.054 (0.05)	0.060 (0.01)	0.023 (0.94)
Age-squared[a]	–0.0036 (0.27)	–0.00023 (0.40)	–0.00057 (< 0.001)	–0.0075 (0.38)	–0.0007 (0.08)	–0.00087 (0.01)	0.00015 (0.97)
Non-white	–0.013 (0.75)	–0.036 (0.31)	–0.063 (0.005)	–1.40 (0.21)	–0.077 (0.15)	–0.11 (0.02)	0.0051 (0.99)
Smoking							
< 10/day	–0.0085 (0.86)	–0.033 (0.42)	–0.038 (0.14)	1.93 (0.13)	–0.024 (0.70)	–0.057 (0.28)	0.94 (0.16)
> 10/day	–0.064 (0.31)	–0.064 (0.23)	–0.082 (0.01)	1.23 (0.47)	–0.14 (0.08)	–0.14 (0.05)	0.063 (0.94)
Recent fever	–0.17 (0.04)	–0.15 (0.03)	–0.021 (0.64)	–7.04 (0.001)	–0.23 (0.03)	–0.25 (0.005)	–1.45 (0.21)
History of STD	–0.088 (0.06)	–0.074 (0.06)	–0.026 (0.29)	–2.12 (0.09)	–0.16 (0.006)	–0.12 (0.02)	–1.62 (0.01)
Ejaculation abstinence time (hr)	0.0024 (< 0.001)	0.0024 (< 0.001)	0.0011 (< 0.001)	0.0078 (0.42)	0.0039 (< 0.001)	0.0036 (< 0.001)	–0.0050 (0.12)
Time to start of semen analysis (hr)	0.082 (0.32)	0.0050 (0.91)	0.022 (0.37)	–4.20 (0.005)	0.068 (0.52)	–0.0034 (0.95)	–0.43 (0.38)
Time to conduct semen analysis (hr)	0.064 (0.35)	–0.016 (0.53)	–0.0064 (0.66)	–2.5 (0.006)	–0.030 (0.73)	–0.049 (0.15)	–0.51 (0.08)

Values shown are regression coefficients (*p*-values) from mixed models adjusted for center and all variables in Table 3.
[a]Using mean values for all other variables, the adjusted TMC (using µ-Cell) for a white man in Minnesota is 113 × 10^6 at age 25 years versus 108 × 10^6 at age 45 years.

(*p*-value < 0.001 for both age and age-squared). Non-whites had significantly lower semen volume than whites. Smoking more than 10 cigarettes/day was associated with decreased semen volume, but had little effect on concentration, motility, or morphology. Fever within the prior 3 months significantly decreased sperm concentration and motility, but not morphology or semen volume. The percent morphologically normal sperm was reduced among men who reported a history of an STD. TMC was significantly associated with all of these covariates, reflecting their relationship to sperm concentration, percent motile sperm, and semen volume, from which TMC is calculated. Similarly, significant associations were seen between total count and race, age (and age-squared), smoking, fever, and history of STDs, reflecting associations between these covariates and sperm concentration and semen volume (data not shown).

Abstinence time (restricted to 2–240 hr) was strongly and linearly related to sperm concentration and semen volume and TMC (all *p*-values < 0.001). Increasing time from sample collection to start of semen analysis and increasing time to complete the semen analysis were each associated with reduced motility.

Little or no association was seen between any semen parameter and use of steroids, self-reported urogenital abnormalities, or history of infertility, all of which were quite rare in this population. No consistent pattern was found between season and any semen parameter, either overall or within each center.

Sensitivity analyses. For the semen analysis we had excluded 12 men on the basis of abstinence times that were missing (*n* = 5), < 10 min (*n* = 2), or > 10 days (*n* = 5), as well as 7 men whose sperm concentration was more than 3 SD from predicted by the modeled relationship between concentration and abstinence

time. To test the sensitivity of the results to these exclusions, we reran the model including the 11 men with an abstinence time between 30 min and 2,000 hr. Their inclusion did not alter the study's conclusions. In fact, the contrasts between Missouri and both New York and California were somewhat stronger, and *p*-values were unchanged or reduced slightly for all between-center contrasts and all semen parameters.

Because the number of subjects in New York was small, we also reran the model excluding these subjects. The effect was to slightly (2–5%) increase the contrasts between Minnesota and Missouri for all semen parameters, so that the model including New York subjects presented here slightly underestimates these differences.

Analyses to examine selection bias. We examined selection bias in two ways. One was to compare participants (*n* = 514) who gave semen samples with those who did not (*n* = 107) with respect to characteristics related to semen quality and fertility (race, age, education, smoking, recent fever, infertility, STD history, and TTP). There were no statistically significant differences between these groups for any of these factors. We also examined selection bias by comparing responses about fertility (whether either partner ever saw a doctor for infertility and TTP) from a sample of study subjects who completed the mini-questionnaire (*n* = 338) and from potential subjects who refused participation (*n* = 956). These fertility-related responses did not differ significantly between groups, although nonparticipants appeared to have somewhat longer TTP. Together, these analyses argue against significant selection bias in this data set.

Discussion

Our study found significantly lower sperm concentration and TMC in fertile men from mid-Missouri relative to those from New York, Minnesota, and California. The percent of sperm that were motile also varied significantly among centers. Differences in semen volume and percent normal sperm (by strict morphology) were small and nonsignificant.
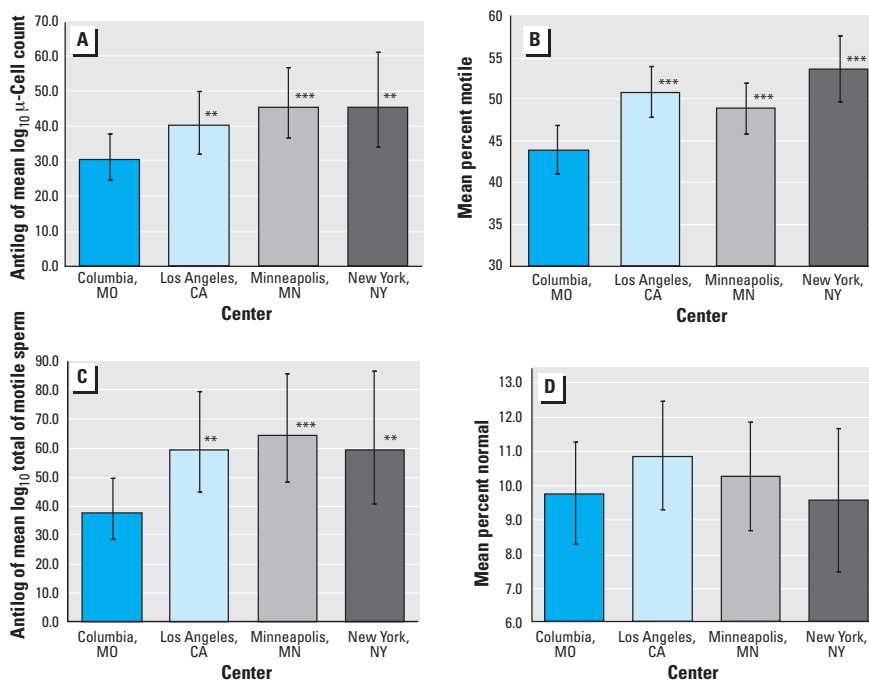

Figure 1. Selected semen parameters by center: (*A*) mean count (µ-Cell); (*B*) percent motile sperm; (*C*) total motile sperm (µ-Cell); and (*D*) percent normal sperm. Error bars indicate 95% confidence limits of the estimated means.
**p* < 0.01, and ***p* < 0.001 as compared to Columbia, MO, using a mixed model adjusting for age, race, smoking, recent fever, history of STD, abstinence time, and analysis time.

Table 4. Summary of adjusted semen characteristics by center.[a]

| | Sperm concentration (10⁶/mL) | | Volume (g) | Percent motile | TMC (10⁶) | | Percent normal sperm |
	Hemacytometer	µ-Cell			Hemacytometer	µ-Cell	
No. samples	466	890	889	891	465	888	876
Center							
Missouri (reference)	35.0	30.8	2.9	43.9	45	37.7	9.8
California	43.4 (0.060)	40.3 (0.005)	3.1 (0.39)	50.9 (< 0.001)	64.6 (0.02)	59.8 (< 0.001)	10.9 (0.11)
Minnesota	54.9 (< 0.001)	45.7 (< 0.001)	3.0 (0.43)	48.9 (< 0.001)	83.2 (< 0.001)	64.5 (< 0.001)	10.3 (0.35)
New York	58.2 (0.001)	46.0 (0.002)	2.5 (0.07)	53.7 (< 0.001)	77.0 (0.007)	59.6 (0.005)	9.6 (0.84)

Values shown are regression coefficients (*p*-values for comparison to Missouri) from mixed models adjusted for covariates listed in Table 3. Analysis was restricted to samples with abstinence times between 2 and 240 hr. Concentration, volume, and TMCs were log-transformed for analysis; estimates are the back-transformed means.
[a]From mixed models including multiple samples per man (except hemacytometer).

The National Cooperative Reproductive Medicine Network, using methods similar to those employed here (Guzick et al. 2001), classified men into three categories: fertile, subfertile, and of uncertain fertility. Men classified as fertile were those with sperm concentration (using µ-Cell) that exceeded 48×10^6/mL with > 63% motile sperm and > 12% morphologically normal sperm. In our study of fertile men, there were significantly fewer samples from men living in mid-Missouri that met all three of these criteria compared with men in the three urban centers (1.1% compared with 8.5%, $p < 0.001$).

We examined three contrasts among the four centers (Missouri–Minnesota, Missouri–California, and Minnesota–New York). Of these, the Missouri–Minnesota contrast is least likely to be affected by confounding and selection bias. The recruitment rates at these two centers were comparable, and their study populations were quite similar demographically. It is reassuring, therefore, that differences in semen parameters between Missouri and Minnesota were large and highly significant.

Although some confounding may remain uncontrolled, we feel this is unlikely to explain the between-center differences we report here. We examined methodological variables (abstinence time, time to start, and complete semen analysis) and adjusted for these. Several personal characteristics of the men were related to semen quality and varied across centers (age, race, smoking, history of STD, and recent fever). After statistical adjustment for these factors, estimates of between-center differences were similar to (or slightly greater than) unadjusted estimates.

In addition, these findings are not likely because of differences in study methods at the four study centers. Common protocols and study instruments were used at all centers. All andrology technicians were centrally trained, and equipment and supplies were standardized across centers. Moreover, strict quality control procedures were implemented, and quarterly testing was conducted throughout the period of the study.

Our study was designed in collaboration with the International Study of Semen Quality in Partners of Pregnant Women (Jorgenson 2001), and protocols and quality control samples were shared between the two studies. How do results of these two studies compare? Differences in concentration in the European study (Jorgenson 2001) were somewhat less marked than those we report here. For example, sperm concentration (by hemacytometer) in Copenhagen (whether using means or medians, adjusted or unadjusted) was 74% of that in Turku (Jorgenson 2001). In comparison, sperm concentration (by hemacytometer) in Missouri was 57% of that

New York and 60% that of Minnesota. Between-center difference in the European study increased somewhat after statistical adjustment, whereas semen volume and sperm morphology varied little among study locations. The four European centers were in urban areas (Copenhagen, Denmark; Paris, France; Edinburgh, Scotland; and Turku, Finland). Although use of agricultural chemicals may differ among these urban centers, these agents have not yet been examined in relation to semen quality.

Using data from prevasectomy males, Fisch and Goluboff (1996) reported mean sperm concentrations of 132, 101, and 73×10^6/mL in New York, Minnesota, and California, respectively. In our study, the urban centers also differed somewhat among themselves, but less than each differed from mid-Missouri. We saw lower sperm concentrations in California than New York and Minnesota, as did Fisch and Goluboff, but, unlike that study, we saw little or no difference in semen quality between Minnesota and New York.

Most studies of semen quality have been conducted in large metropolitan areas, and it is difficult to find comparable studies from semi-rural areas. Among the 61 studies analyzed by Carlsen et al. (1992) in a much-cited meta-analysis, 27 were conducted in the United States. Of these, only one, in Iowa City, Iowa, was conducted in a county of < 250,000 residents (Nelson and Bunge 1974). In this Iowa population, the mean (hemacytometer) sperm concentration in prevasectomy patients was 48×10^6/mL, which is lower than the concentration reported here for Columbia, Missouri. We compared population density, proportion of land in farms, and use of agricultural chemicals for the four centers in the current study as well as Iowa City, Iowa (U.S. Census Bureau 2001). Acres in farmland ranged from 288,139 in Johnson County (where Iowa City is located) and 249,849 in Boone County (where Columbia, MO, is located) to 69,128 in Minnesota and 0 in New York City. Agricultural chemicals (fertilizers, pesticides, or herbicides) were applied to all (or most) of this farmland. A recent U.S. Geological Survey report on water quality (U.S. Geological Survey 2001) noted that extensive herbicide use in agricultural areas (accounting for about 70% of total national use of pesticides) has resulted in widespread occurrence of herbicides in agricultural streams and shallow groundwater in those areas. We are examining urinary metabolite levels in relation to semen quality in a subset of the population in a separate analysis and hope to obtain funding to obtain biomarkers of pesticide exposure on the entire study population. When data from the entire cohort have been collected, we will examine semen quality with respect to self-reported pesticide exposures as well.

This study has a number of strengths but also some weaknesses. Among its strengths are its prospective design and strict adherence to protocol to ensure comparability across centers. The exacting quality control demands, for all aspects of the study, have produced semen analysis results of extremely high precision.

As with all studies of semen quality, low participation rates and potential selection bias are of concern. In studies of partners of pregnant women, recruitment is particularly difficult because a woman must give permission before her partner can be contacted, unless he is present at the prenatal visit. To examine selection bias, we compared questionnaire data on TTP and history of infertility, as well as demographics, of study subjects and nonparticipants, and of men who did and did not give semen samples. Reassuringly, there was little evidence that these populations differed. However, the limited number of non-white subjects, and few subjects from New York limited our ability to examine this question within ethnic groups and at all centers.

The current analysis is not able to explain the between-center difference in semen quality demonstrated. However, the extensive data (questionnaire and biological samples) available on these men will permit us to examine a range of hypotheses in future analyses.

In the current study we found considerably reduced semen quality in Columbia, Missouri, compared with New York, Minnesota, and California. Although there may well be multiple factors on which Missouri differs from the other centers, Missouri is unusual among sites for semen studies because of its proximity to intensive agriculture. The limited availability of semen quality data from semirural, agricultural communities, the historically low concentrations in Iowa, and the low sperm concentration and percent motile sperm reported here for Columbia, Missouri, suggest the need for further study in such communities.

REFERENCES

Auger J, Jouannet P. 1997. 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 12:740–745.

Berman NG, Wang C, Paulsen CA. 1996. Methodological issues in the analysis of human sperm concentration data. J Androl 17:68–73.

Brouwer A, Ahlborg UG, van L Feeley MM. 1998. Report of the WHO working group on the assessment of health risks for human infants from exposure to PCDDs, PCDFs and PCBs. Chemosphere 37:1627–1643.

Carlsen E, Giwercman A, Keiding N, Skakkebaek NE. 1992. Evidence for decreasing quality of semen during past 50 years. Br Med J 305:609–613.

Fenster L, Katz DF, Wyrobek AJ, Pieper C, Rempel DM, Oman D, et al. 1997. Effects of psychological stress on human semen quality. J Androl 18(2):194–202.

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

Guzick DS, Overstreet JW, Factor-Litvak P, Brazil CK, Nakajima ST, Coutifaris C, et al. 2001. Sperm morphology, motility, and concentration in fertile and infertile men. N Engl J Med 345:1388–1393.

Jorgenson N, Anderson A-G, Eustache F, Irvine DS, Suominen J, Petersen JH, et al. 2001. Regional differences in semen quality in Europe. Hum Reprod 16: 1012–1019.

Laird NM, Ware JH. 1982. Random-effects models for longitudinal data. Biometrics 38:963–974.

Nelson CM, Bunge RG. 1974. Semen analysis: evidence for changing parameters of male fertility potential. Fertil Steril 25:503–507.

Overstreet JW, Brazil C. 1997. Semen analysis. In: Infertility in the Male (Lipshultz LI, Howards SS, eds). 3rd ed. St. Louis, MO:Mosby, 487–490.

Paulsen CA, Berman NG, Wang C. 1996. 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 65:1015–1020.

SAS Institute. 2001. The SAS System for Windows, Version 8.2. Cary, NC:SAS Institute, Inc.

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

U.S. Census Bureau. 2001. State and County QuickFacts. Washington, DC:U.S. Department of Commerce, U.S. Census Bureau. Available: http://quickfacts.census.gov/qfd/ [accessed 28 March 2002].

United States Geological Survey. 2001. Water-Quality Patterns in Agricultural Areas. Reston, VA:U.S. Department of the Interior, United States Geological Survey. Available: http://water.usgs.gov/pubs/circ/circ1225/html/wq_agric.html [accessed 7 August 2002].

WHO Task Force on Methods of Regulation of Male Fertility. 1996. Contraceptive efficacy of testosterone-induced azoospermia and oligozoospermia in normal men. Fertil Steril 65:821–829.

Wittmaack FM, Shapiro SS. 1992. Longitudinal study of semen quality in Wisconsin men over one decade. Wis Med J 91:477–479.

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

———. 1999. WHO Laboratory Manual for the Examination of Human Semen and Sperm-Cervical Mucus Interaction. 4th ed. Cambridge, UK/New York:Cambridge University Press.

Zeger SL, Liang KY. 1986. Longitudinal data analysis for discrete and continuous outcomes. Biometrics 42:121–130.