Pure Appl. Chem., Vol. 75, Nos. 11–12, pp. 2181–2193, 2003.

Topic 3.16
Environmental estrogens and sperm counts*

Harry Fisch‡ and Robert Golden

Department of Urology, Columbia University, 944 Park Ave., New York, NY 10028, USA

Abstract: The term "environmental estrogen" refers to chemical substances that exhibit some degree of estrogen-like activity. The primary emphasis for potential adverse effects resulting from exposure to environmental estrogens is on in utero exposure because such exposure can occur during critical periods of organogenesis. Assessment of biological plausibility can be based, in part, on the extensive data on the effects of diethylstilbestrol (DES). The available evidence is too limited to judge with any confidence whether sperm counts have declined during the past 50 years. Based on both animal and human data with DES, it is biologically plausible that in utero exposures to exogenous estrogenic compounds are capable of reducing sperm production in adult men. However, the apparent existence of a maternal dose threshold for DES-induced effects on sperm counts undermines the likelihood that environmental estrogens, which are substantially less potent, are capable of causing similar effects.

INTRODUCTION

There is currently great uncertainty whether some synthetic chemicals, acting as "hormone-mimics" and released into the environment, might have the potential to disrupt the endocrine systems of humans. Although there is no question that hormones are potent modulators of biochemical and physiological function, the implication that in utero exposures to environmental endocrine modulators (e.g., xenoestrogens) have the capability to produce clinically detectable effects in humans is much less certain.

Although exogenous endocrine modulators can involve any hormonal system, the principle focus of this review is on estrogen-mediated effects. As used in this review, the term "environmental estrogens" refers to chemical substances that exhibit some degree of estrogen-like activity. In order to be accurate, this term must apply to both naturally occurring estrogenic compounds (e.g., phytoestrogens) and to synthetic estrogenic compounds. Naturally occurring estrogens in either humans or animals are referred to as endogenous estrogens, and all compounds with estrogenic properties entering the body from an outside source are referred to as exogenous estrogens.

The primary emphasis here for potential adverse effects resulting from exposure to environmental estrogens is on in utero exposure because such exposures can occur during critical periods of organogenesis. The concept of a critical window of exposure or sensitivity is an important factor that must be considered in animal studies and in human exposure situations. Abundant animal and human data demonstrate this phenomenon. For some adverse health outcomes (e.g., reproductive tract abnormalities) exposures to excessive amounts of estrogen during this critical period is essentially the only way that such effects could occur. For other adverse health endpoints, exposure to estrogen in utero may be a risk factor for subsequent adverse effects.

---

*Report from a SCOPE/IUPAC project: Implication of Endocrine Active Substances for Human and Wildlife (J. Miyamoto and J. Burger, editors). Other reports are published in this issue, *Pure Appl. Chem.* **75**, 1617–2615 (2003).

‡Corresponding author

POTENTIAL EFFECTS OF CONCERN

The following adverse health consequences to humans have been hypothesized to be associated with in utero exposure to environmental endocrine modulators (primarily environmental estrogens) [1,2]:

- adverse effects on male reproductive tract
- male and female fertility problems
- testicular cancer
- prostate cancer
- breast cancer
- endometriosis
- learning disability or delay
- alterations of sexual behavior
- immune system effects
- thyroid effects

It is biologically plausible to hypothesize that exposure (particularly in utero) to environmental estrogens could adversely affect humans since exposure to estrogen, whether in utero or in adulthood, can have biochemical, physiological, and specific target organ effects on development, reproduction, behavior, and metabolism [3,4]. Consequently, it is not unreasonable to hypothesize that exposure to estrogen-like compounds, whatever their source, could adversely affect human health. However, biological plausibility alone is not a sufficient basis for concluding that exposure to compounds identified as environmental estrogens has adversely affected humans.

CRITERIA FOR ASSESSING BIOLOGIC PLAUSIBILITY

DES paradigm

In assessing the likelihood that in utero exposure to environmental estrogens might be etiologically linked to any particular adverse effect, it is important to consider biological plausibility as part of the evaluation process. Assessment of biological plausibility can be based, in part, upon the extensive data on the effects of diethylstilbestrol (DES). There are considerable data, in both experimental animals and humans, on the effects of in utero exposure to DES on adult offspring. While the DES data can be used to judge the likelihood of adverse effects resulting from in utero exposure to environmental estrogens, because of the manner in which DES was used clinically, these data are not useful for assessing the likelihood of adverse effects that might result from postnatal exposure to environmental estrogens except in the mothers. Most of the adverse effects observed in both animals and humans following in utero exposure to DES are consistent with basic principles of dose response, as well as with the presence of maternal dose levels not associated with adverse effects in offspring. As reviewed in greater detail below, the clinical use of DES resulted in the massive exposure to 4 to 6 million women and their unborn children. Critical reliance upon the abundant DES database provides a useful tool for assessing many of the endpoints of concern according to the following scenarios:

- If an effect hypothesized as possibly resulting from in utero exposure to environmental estrogen is not reliably associated with in utero exposure to DES, it may be biologically implausible that such an effect (even if real) is mediated via an estrogen mechanism of action.
- If an effect hypothesized as possibly resulting from in utero exposure to environmental estrogen is similar to an effect reported following in utero exposure to DES, the likelihood that such an effect might be associated with exposure to environmental estrogen must be evaluated based upon comparative potency and dose–response considerations.

For some of the specific effects addressed in this review, it is imperative to note that the weight of the available evidence currently may be insufficient to reach any valid conclusion. The data may sim-

ply be inadequate to conclude one way or another whether in utero exposure to exogenous estrogen might be a risk factor.

DES as a model compound to judge potential effects of exogenous estrogen

DES is a synthetic estrogenic compound that is several times more potent than 17β-estradiol and hundreds to thousands of times more potent (based primarily on results from in vitro tests) than other compounds identified as exogenous environmental estrogens (not including oral contraceptives or estrogen replacement therapy), whether synthetic organochlorine compounds [e.g., certain metabolites of dichlorodiphenyltrichloroethane (DDT), selected polychlorinated biphenyl (PCB) congeners, aldrin, dieldrin, endrin] or naturally occurring phytoestrogens from plants [5–9]. The large difference in potency between DES and environmental estrogens may be important in assessing the likelihood that in utero exposure to environmental estrogens might be a risk factor for certain adverse health endpoints. Based upon substantial differences in estrogenic potency, the total maternal dose of DES required to produce adverse effects in offspring exposed in utero and maternal intakes of exogenous environmental estrogens that might be required to result in comparable estrogenic activity can be compared. While such comparisons may provide only a rough approximation, they can highlight potentially important qualitative (and perhaps even quantitative) differences between the dose–response characteristics of different estrogen receptor agonists.

It is critical to note that the approach outlined above is based on an estrogenic potency comparison between DES and various individual compounds identified as environmental estrogens. This, of course, is not representative of how real-world exposures to these compounds occurs. In reality, typical exposures, whether as a result of stored body burdens or from daily dietary intake, are far more likely to be to a complex mixture of weakly estrogenic compounds. Consequently, even though individual environmental estrogens may be less potent than natural estrogen or DES, there are two critical questions at the core of this issue that must be addressed. First, what are the *net* effects of simultaneous exposure to many such "weak" estrogens? Second, will the estrogenic sum of all such chemicals acting together at the level of the estrogen receptor result in effects that would not occur from exposure to just a few of these compounds? Speculation has centered around the idea that background adipose tissue concentrations and resulting serum levels of a number of persistent, bioaccumulating, weakly estrogenic compounds might have additive (or synergistic) net estrogenic activity capable of producing adverse effects. Several studies have attempted to address this issue. For example, when 10 different chemicals [i.e., endosulfan α and β, toxaphene, dieldrin, 2,3,4,5-tetrachlorobiphenyl, *p,p′*-DDT, 2,2′,3,3′,6,6′-hexachlorobiphenyl, *p,p′*-dichlorodiphenyldichloroethylene (*p,p′*-DDD), *p,p′*-dichlorodiphenyldichloroethylene (*p,p′*-DDE), and methoxychlor] already known to be weakly estrogenic in vitro in 7 (MCF-7) cells, were added to MCF-7 cells at one-tenth the concentration required for each to produce an estrogenic effect, the results indicated additive net estrogenic activity similar to 17β-estradiol [8].

In a provocative in vitro study using yeast cells containing human estrogen receptor, selected chemicals (i.e., endosulfan, dieldrin, toxaphene, and chlordane) were tested for estrogenic activity singly and in various combinations [10]. The results of this study suggested that endosulfan, dieldrin, and toxaphene had very low estrogenic potency, while chlordane was devoid of any estrogenic activity. However, various binary combinations of these chemicals demonstrated estrogenic potency from 160 to 1600 times greater than the activity of any single chemical. Chlordane, which had no estrogenic activity when tested alone, significantly enhanced the potency of the other chemicals tested. In addition, two PCB compounds were also tested in the same system (2′,4′,6′-trichloro-4-biphenylol and 2′,3′,4′,5′-tetrachloro-4-biphenylol) and shown to exhibit a similar pattern of synergistic estrogenic activity. For the two PCBs tested, the result in yeast cells was confirmed in Ishikawa cells, an endometrial cancer cell line transfected with human estrogen receptor. However, numerous attempts to replicate these findings have not been successful. Ten different estrogen-responsive assays (including induction

of uterine wet weight, peroxidase activity, and progesterone receptor levels; induction of cell growth in MCF and human breast cancer cells; induction of reporter gene activity in two yeast-based assays; and competitive binding to human and mouse estrogen receptor) were used to test all of the binary mixtures noted above. The results confirmed additive activity, but no synergism was observed [11]. In two additional recent studies, one on a mixture of dieldrin and endosulfan tested in a yeast human estrogen transactivation assay and a rat uterotropic assay [12], and another on a dieldrin-toxaphene mixture tested in a mouse uterotropic assay, MCF-7 human breast cancer cells, and a yeast-based reporter gene assay, also failed to demonstrate any evidence of synergism [13]. Finally, the results of the original study reporting synergism [10] were recently withdrawn by the authors [14].

Effects of DES in humans

While estimates vary, DES was widely prescribed to several million pregnant women, with most use occurring between the late 1930s until 1971, in the mistaken belief that it was useful in the treatment of threatened or habitual abortion [15–17]. The use of DES during pregnancy was proscribed by the Food and Drug Administration (FDA) in 1971 with the discovery that a small number of women exposed in utero later developed vaginal clear cell adenocarcinoma [17,18]. As reviewed below, in utero exposure to sufficient maternal doses of DES is also associated with other adverse effects on the reproductive tracts of both males and females. In males, these effects include epididymal cysts, microphallus, cryptorchidism, testicular hypoplasia, decreased sperm count, and increased incidence of abnormal sperm [16,19–21]. In females, reported adverse effects include vaginal adenosis; clear cell adenocarcinoma; structural defects of the cervix, vagina, uterus, and fallopian tubes; infertility; and irregular menstrual cycles [16,19,22–24].

Because of the widespread clinical use of DES, there are some dose–response data on a number of adverse effects in both male and female offspring exposed in utero. During the peak years of DES use, different DES dosing regimens were used in different medical centers. Since the efficacy of DES had never been established by adequate clinical testing, there was no standard dosing regimen in use. This resulted in large groups of women receiving substantially different total DES doses during pregnancy. Such data permit rough approximations of the daily doses and total maternal doses of DES associated with some of the adverse effects on offspring exposed in utero. Table 1 summarizes the existing data on DES cohorts.

Table 1 Summary data on DES-exposed cohorts and estimated maternal exposures.

Cohort	Type of study	Estimated mean	Persons exposed total DES dose (g)	Non-DES-exposed controls
Mayo Clinic Sons [25]	Cohort	1.4	828	676
Connecticut Mothers [26]	Cohort	2.1	1531	1404
DESAD Study [27]	Cohort	4.2[a]	4014	24
Boston Collaborative Study [28][b]	Cohort	6.4	217	1033
DES Efficacy Trial [15,29,30][c]	Randomized clinical trial	11.6	840	
British Radomized Trial [31]	Randomized clinical trial	11.5	650	806
British Medical Research Council [32]	Randomized clinical trial	17.9	70	66

[a]Median dose based on 26 % of cohort with known DES doses.
[b]80 % of cohort from Boston Lying-In Hospital; DES mean total dose estimated for total cohort.
[c]University of Chicago.

Of central interest to the subject matter of this review is the fact that there are clinical studies that appear to identify no-effect total maternal dose levels and durations of exposure to DES for some adverse effects. These studies suggest that certain maternal DES dosing regimens were not sufficient to result in certain adverse effects to offspring exposed in utero. For example, a study conducted on a cohort from the Mayo Clinic, which appears to have employed the lowest DES dosing schedule of any clinical center in the United States, demonstrates a mean no-effect maternal dose level for effects attributable to in utero DES exposure [25]. This study was designed to determine whether males exposed in utero to DES had a higher frequency of urogenital abnormalities than an unexposed control group. Comparison of 265 DES-exposed males with 274 controls revealed that in utero exposure to DES did not increase the risk of any urogenital abnormalities, including penile length or diameter, testis length or width, epididymal cysts, or varicocele. There were also no adverse effects on sperm volume, density, motility, morphology, or fertility. The mean total administered maternal dose of DES was 1.4 g, with a median duration of exposure of 101 days. While it is possible that more subtle effects might have been present, only the effects noted above were studied. A potential problem with any such studies (whether in humans or animals) is that fetal exposure to DES may have occurred after the critical period of reproductive tract development. Records of DES doses and durations of exposure in this cohort reveal that the 25^{th} and 50^{th} percentiles of the first gestational day of exposure were on days 54 and 89, with exposure durations of 28 and 100 days, respectively [25]. Therefore, the timing of DES exposure is well within the window during which development of the reproductive tract is occurring. While the correlation between some endpoints (e.g., adenosis) and maternal doses is not perfect, the clinical data can provide an approximate benchmark level for investigating potential relationships between maternal exposure to DES (and by implication to other less potent environmental estrogens) and adverse effects in offspring exposed in utero.

ADVERSE EFFECTS ON MALE REPRODUCTIVE TRACT

Decreased sperm quality

Introduction

A great deal of attention has recently been directed to the allegation that sperm quality has been declining worldwide during the past 40 to 50 years [33–35]. Much of the attention in this debate has focused on the hypothesis that in utero exposure to environmental estrogens might be responsible. While certain chemicals have the ability to affect sperm quality if exposure to sufficient doses occurs over a sufficient period of time, such effects must be evaluated on a compound-by-compound basis. The questions raised with respect to the issue of declining sperm quality are extraordinarily complex, and much of the data collected to date are conflicting. Because of the profound implication of the claim that sperm quality (particularly sperm counts) worldwide has declined over the past 40 to 50 years, the evidence must be critically evaluated. Either adult or prenatal exposure to certain chemicals (both estrogenic and nonestrogenic chemicals) can affect sperm quality (i.e., counts, motility, or morphology). However, before it can be concluded that chemicals, whether estrogenic on nonestrogenic, have the potential to affect sperm quality, particularly worldwide, it is important to assess dose response and exposure data, potency, and possible interactive effects.

It is biologically plausible that in utero exposure to estrogen can affect sperm counts. In studies with mice, in utero exposure to DES is associated with decreased sperm production and abnormal sperm morphology in adult offspring [6,36]. There are also numerous studies of adult male humans prenatally exposed to DES; however, with respect to decreased sperm counts, the results of these studies are mixed. Decreased sperm count (115 million/ml in 87 controls vs. 91 million/ml in 134 DES-exposed) and abnormal Eliasson scores in 18 % of 134 DES-exposed men compared to 8 % in 87 placebo-exposed controls have been reported [29]. This study was conducted on a cohort of men exposed in utero to DES according to the dosing protocol in use at the University of Chicago where mean total maternal

DES dosages were, on average, 11 603 mg. Another small study reported lower average sperm counts in 20 DES-exposed men and pathologic Eliasson scores in 18 DES-exposed men, but no comparison controls were used and maternal DES dosages were unknown [20]. In contrast, in a cohort of men from the Mayo Clinic exposed in utero to DES in which total mean maternal DES doses were approximately 1.4 g, there were no differences between 110 DES-exposed men and unexposed men in sperm count, motility, or abnormal Eliasson scores [25].

While clinical data suggest that in utero exposure to DES at some maternal dose level can result in decreased sperm counts, the data also suggest the existence of an apparent maternal dose threshold for such effects. Even at the dose levels of DES that produced adverse effects on sperm following in utero exposure, the magnitude of such effects is not large, with average sperm counts reduced from 115 million/ml in men not exposed to DES to 91 million/ml in DES exposed men. While this shows a DES effect, it also demonstrates that an estrogen as potent as DES does not produce a decrease in mean sperm counts approaching the levels at which fertility might be affected. The failure of the maternal DES dosages used at the Mayo Clinic to cause any effects on sperm count or morphology underscores this point. As noted later in this review, fertility in the high-dose cohort of men (University of Chicago cohort) was not adversely affected, even after in utero exposure to DES at maternal doses sufficient to produce adverse effects on sperm counts and morphology. This raises the question of whether in utero exposure to environmental estrogens, which are significantly less potent than DES, could cause a decline in sperm counts.

Evidence for declining sperm quality

The critical evidence cited in support of the hypothesis that sperm quality has declined over the past 50 years is the Carlsen et al. [33]. Regression meta-analysis of 61 studies showing a steady decline in mean sperm counts from 1938 to 1990. The results of this study have been interpreted as demonstrating that in utero exposure to compounds identified as environmental estrogens (e.g., DDT, PCBs, and other hormonally active chemicals) might be responsible for this effect. Because of the biologic plausibility that prenatal exposure to estrogenic substances could affect adult sperm counts, it is important to determine whether decreased sperm counts are a real phenomenon, and if so, whether decreases are caused by exposure to environmental estrogens.

In general, the meta-analysis relied upon 61 studies selected from the years 1938 to 1990 that reported mean sperm counts; results were weighted by the number of subjects in each study [33]. While this approach appears valid, closer scrutiny of the methodology and data reveal a number of potential problems. It may not be reasonable to assume that methods and protocols for sperm collection and measurement from laboratories all over the world from 1938 to 1990 would be similar. Even in the United States, there is little standardization of methods for semen analysis, which suggests a substantial potential for error when studies spanning more than 50 years from all over the world are compared. There is also a wide range of biological variability in sperm counts, as well as numerous factors that may affect results, including temperature, season, period of abstinence, and others. Even though Carlsen et al. excluded studies that reported counting with computer-assisted or flow cytometric methods, the methodology relied upon (i.e., the use of various types of counting chambers) is fraught with numerous potential interpretive problems [33]. Depending solely on the semen counting chamber used (i.e., Cell VU, Microcell, Neubauer Hemacytometer, or Makler) mean sperm counts can vary from 35 to 51 million/ml based on calibration with glass beads of known size and concentration. Other documented technical sources of error include inter-individual counting differences, counter chamber variation, and poor pipetting technique [37]. Other studies demonstrate similar variation in sperm counts based on the counting chamber used. In particular, the use of the Makler chamber results in counts that can be as much as 65 % higher with significantly higher standard deviation, suggesting that counts made with this device are less reproducible [38]. There is also great variability in individual sperm counts from one ejaculate to the next, with counts varying from <20 million/ml to >100 mil-

lion/ml [37]. Daily sperm production is also significantly reduced in aging men as is the number of Sertoli cells, which is also correlated with daily sperm production. Vascular degeneration, autoimmunity, or loss of Sertoli cells may contribute to the age-related decline in sperm production [39]. In evaluating the 61 studies in the meta-analysis, it is not clear to what extent age has been considered as a possible confounding factor. Finally, temperature has a profound effect on semen quality. In a pairwise study of 131 men who contributed semen specimens in summer and winter, there were significant decreases during the summer in sperm concentration, total sperm count per ejaculate, and concentration of motile sperm. Mean reductions for these parameters were 32, 24, and 28 %, respectively. These reductions correlated with lower birth rates during spring months [40]. These data demonstrate that seasonal variations in sperm counts may influence comparisons that fail to account for this phenomenon. The various factors noted above may be problems when sperm count results from different laboratories or clinics, in which different counting methods were used, are compared. In addition, certain methodologic practices may introduce systematic counting errors that would tend to produce results consistent with an apparent effect, when, in reality, no such effect was occurring. Significant intra- and interlaboratory variation also make it difficult to make meaningful comparisons of sperm morphology over time. These issues will require careful analysis before the results of comparative studies can be accepted as demonstrating a real effect.

Data not included in the 1994 meta-analysis

Additional studies that appear to refute the hypothesis

Additional studies, not included in the 1992 meta-analysis, appear to refute the hypothesis of declining sperm counts. The outcomes of these studies, which have not been integrated into the linear regression curve, illustrate the sensitivity of conclusions to data collection efforts. For example, for 1984, five studies were cited on a total of 290 men with mean sperm concentrations [reported in million/ml; (number of men)] of 83.9 (119), 72 (114), 58.9 (9), 59 (36), and 102 (12). Not cited was another 1984 study that reported mean sperm concentrations of 107 million/ml on 861 men [41]. This 1984 mean sperm concentration value, based on more subjects than 12 of the 13 pre-1970 studies relied upon, does not appear to support the hypothesis that sperm counts have declined. The meta-analysis precedes a 1992 study that reports mean sperm concentrations of 101.5 million/ml in 30 men [42], a level of sperm density similar to that in 1950. A longitudinal study of semen quality in Wisconsin men, also published in 1992, showed no decline in sperm concentration, motility, or ejaculate volume between 1978 and 1987. Mean sperm counts ranged from 55.6 million/ml to 105.7 million/ml [43]. A sharp rise in morphologically abnormal sperm between 1982 and 1983 coincided with a change in criteria used to identify abnormal sperm. A detailed historical analysis of sperm counts from 1951 to 1977 concluded that there was no significant change during this time period [44]. This interval, which encompasses about half of the 50-year period reviewed by Carlsen et al. [33], appears to refute the hypothesis that sperm counts have fallen over the years. Additional data are provided by a study of in utero exposure to maternal tobacco smoke and decreased semen quality in male offspring [45]. In this study, the mean sperm concentration over all exposure groups (n = 292) was approximately 106 million/ml. These data appear to have been collected in the United States in 1991. There is a marked contrast between the mean sperm count from this study and the mean sperm count of 54 million/ml (n = 54) reported by Carlsen et al. [33] for 1990, the last year considered in their review. A mean sperm count of 106 million/ml is more in line with the mean sperm count from the meta-analysis regression curve for about 1950 and does not support the hypothesis of declining sperm counts over the past 40 to 50 years. The issue of changes in semen quality over time was investigated in a retrospective analysis of 510 healthy men in the Seattle area who donated multiple semen samples between 1972 and 1993 [46]. The results of this study showed no decrease in sperm concentration, semen volume, total number of sperm per ejaculate, and percent normal sperm morphology over the 21-year period studied. There was actually a small, but statistically significant, increase in these parameters with time, although this finding was not considered

to be biologically important. Multiple semen samples from each participant permitted the use of geometric mean values for each individual, providing more power to detect changes and reduce intrasubject variability due to different abstinence intervals.

In the most comprehensive study to date, Fisch et al. [47] investigated semen quality in 1283 men from the United States over a 25-year period between 1970 and 1994. This study included all men who banked sperm prior to vasectomy at sperm banks in New York, Minnesota, and California (i.e., no man was excluded from analysis with arbitrary rejection criteria). After controlling for age and duration of abstinence, there was a slight, but statistically significant increase in mean sperm concentration [77 million/ml in 1970 vs. 89 million/ml in 1994 ($p < 0.05$)], but no change in motility or semen volume during the 25-year period. A rather dramatic finding was the significant differences in mean sperm concentration and motilities between different sperm banks, with California the lowest (72.7 million/ml), Minnesota intermediate (100.8 million/ml), and New York the highest (131.5 million/ml). This striking geographical variation highlights a potentially significant confounder of studies such as Carlsen et al. [33], which are unable to account for this variable. As noted by Fisch et al. [47], the marked geographic variation found in their study had been reported previously [44,48]. After reviewing the papers relied upon by Carlsen et al. [33] in the meta-analysis, it was determined that the preponderance of earlier studies (which defined the upper end of the regression curve) were from New York, while, after 1970, only 50 % of men in the meta-analysis were from the United States, with only 25 % of these from New York. Based on the demonstrated geographic variability in semen quality across the U.S. and the possibility of similar variability around the world, the meta-analysis of Carlsen et al. [33] may be confounded by the inability to account for geographic variability in the parameters of concern. The likelihood that this has occurred is suggested by a reanalysis of all 61 studies included in the regression meta-analysis [49]. Of the studies comprising the meta-analysis, only 20 included more than 100 men, but these same studies represented 91 % of the total men studied. Of the studies prior to 1970, all were from the United States, with 80 % of these from New York, where sperm counts (then and now) are the highest. After 1970, only three studies were from the U.S., with many from third world countries, where sperm counts were low. Additional conformation of geographic variation is provided by the study reviewed above on 510 men from the Seattle area [46]. Mean sperm counts observed in this study (46.5 million/ml in 1972 and 52 million/ml in 1993) are similar to, but still lower than, mean sperm counts reported for California of 72.7 million/ml, suggesting not only an east–west gradation, but a north–south gradation, as well. Obviously, the significant influence of geographic variation needs to be considered when studies from all over the world are analyzed and compared for potential trends in sperm count over time.

It is beyond the scope of this review to resolve this complex issue. However, it is clear that additional study and analysis remain to be done in order to determine whether the trend reported by Carlsen et al. [33] as well as others is real or an artifact. While it may be tempting to hypothesize that exposure to compounds identified as environmental estrogens might be responsible for the observed effects, it is more likely premature to conclude that sperm counts are declining, much less to ascribe causes.

Alternative statistical analysis of the data

Of the 61 studies in the Carlsen et al. [33] meta-analysis, only 13 are from before 1970, and the results from these studies define the upper end of the regression curve. When only studies from 1970 to 1990 are considered, there is no downward trend in sperm counts. In a reanalysis of the 48 studies published since 1970, sperm counts actually increased between 1970 and 1990 [50]. Another reanalysis suggested that the reported decline in sperm concentration may have been accounted for entirely or in part by a change in lower reference values. This analysis concluded that the hypothesis that sperm counts have declined significantly between 1940 and 1990 was not supportable by the original data [51]. In the most comprehensive statistical reanalysis of the Carlsen et al. [33] study, in addition to reproducing the orig-

inal linear regression, several other statistical models (quadratic, spline fit, and stairstep) were employed to determine the best model to analyze the data [52]. Part of the problem with the data concerns the highly nonuniform distribution of subjects and sperm measurements over the period of time under study (1938 to 1990; the majority of data were collected between 1970 and 1990). Of the four models used, only the linear model was consistent with the hypothesis that sperm counts have declined since 1940. The quadratic and spline fit models suggest that mean sperm counts have been increasing since 1940, while the stairstep model suggests a sharp drop in the mid-1960s, with levels constant since then. There is also an apparent inconsistency between the analysis of Carlsen et al. [33] and the data of MacLeod and Wang [44] noted above. Based on the linear regression model, the last 1000 men in the series studied by MacLeod and Wang [44] should have had a mean sperm count 20 % lower than what was actually observed. Overall, the meta-analysis illustrates the influence of methodologic issues on study results and conclusions. These include potential selection biases, variability in collection and analytical methods, uneven availability of data over the entire time period, and the questionable generalization of data obtained in different geographic areas. It is critical to note that the last 20 years of data, which contain 78.7 % of all the studies and 88.1 % of the total number of subjects, indicate no decrease in sperm counts; in fact, sperm counts were observed to have increased. Finally, a number of other mathematical models, which perform statistically better than the linear model when describing the recent data, suggest a much different explanation of the data. During the last 20 years, the interval with the most robust data, all the models, except the linear model, suggest constant or slightly increasing sperm counts.

Potential causes of observed effects on sperm production

As noted above, unless total in utero exposure to DES exceeded a certain maternal dose level, there do not appear to be adverse effects on sperm production (or any other related parameter) in adult offspring. In light of data suggesting a maternal DES dose threshold level for this effect, the dose–response characteristics and potency aspects of other potential chemical exposures must be considered in any attempts at explaining the observed declines in sperm production, if indeed such declines are real.

Dioxin [e.g., 2,3,7,8-tetrachloro-*p*-dibenzodioxin (TCDD)] is often hypothesized as a potential cause of wide-ranging environmental estrogen effects. With respect to decreased sperm counts, in utero exposure to a single dose of TCDD causes decreased sperm counts in adult male rats [53]. A recent study confirmed that in utero exposure to a single TCDD dose of 1 μg/kg reduces daily sperm production and epididymal sperm reserves [54]. In this study, the authors conclude that the effects of in utero TCDD exposure on sperm production are consistent with the results of Carlsen et al. [33] on declining sperm counts in humans. However, the relevance to humans of this effect in rats must be considered carefully. The single dose of TCDD used to produce a decline in daily sperm output was 1 μg/kg or 1 000 000 pg/kg, administered on day 18 of gestation, a dose level more than six orders of magnitude greater than typical daily intake levels of TCDD (0.3–0.6 pg/kg/day) [55]. Because of the high dose of TCDD used, this study is of questionable relevance for generalizing to human offspring exposed to TCDD in utero. A recent study also calls into question the potential relevance of the above findings. Administration to pregnant rats on day 15 of gestation of a single oral dose of TCDD (0.5, 1.0, 2.0 μg/kg) or indole-3-carbinol (I3C, 1.0 or 100 mg/kg), an Ah receptor agonist found in cruciferous vegetables (i.e., cauliflower, Brussels sprouts, broccoli) produced both common and different reproductive tract abnormalities in adult male offspring exposed in utero. These included decreases in the weights of seminal vesicles, prostate, and testicular parenchyma from one or more doses of TCDD and decreases in daily sperm production from one or more doses of I3C [56]. These findings underscore the complexity of assessing exposure and potential risks, and also (as reviewed in greater detail in the section on behavioral effects) calls into question the relevance of animal models for these kinds of effects.

In contrast to the in utero animal studies, extensive studies of the Air Force Ranch Hand population exposed to Agent Orange (i.e., TCDD) in Vietnam have not demonstrated any significant effects

on sperm count, percent low sperm count, or abnormal sperm among adult men with elevated serum TCDD levels [initially 52–>292 parts per thousand (ppt; median 143 ppt for entire cohort) and currently >33 ppt median levels] [57]. However, studies on the Ranch Hand cohort do not address the issue of possible effects on sperm production following prenatal exposure to TCDD.

Although PCBs have also been hypothesized as potential causes of environmental estrogen effects based on the weak estrogenicity of some congeners, there is, as yet, no evidence suggesting that PCB exposure affects any male reproductive parameters, including sperm counts. A study of 37 currently exposed PCB workers [median serum PCB levels of 12 parts per billion (ppb)] and a comparison group (median PCB levels of 6 ppb), mean sperm concentrations were 65.5 million/ml and 67.2, respectively [58]. No other reproductive abnormalities were observed in this study. As with the Ranch Hand studies, this study also does not address the issue of possible effects subsequent to prenatal exposure to PCBs. Whether massive in utero exposure to PCBs and polychlorinated dibenzofurans (PCDFs) has adverse effects on sperm counts awaits appropriate follow-up studies of the male populations accidently exposed to PCBs and PCDFs at Yusho and Yu-Cheng.

Potential confounders of sperm count studies

Even assuming that sperm counts have declined, attributing any decline to environmental estrogens without consideration of other potential contributing factors may be unwarranted. For example, a significant increase in the use of marijuana worldwide, particularly evident in the 1960s, coincides with reports [33] of declining sperm counts. Several studies have demonstrated that the major active ingredient in marijuana can decrease sperm production and interfere with normal reproductive function and fertility in adult rats following in utero exposure [59–61]. Another possible factor that might play a role is cigarette smoking. There are conflicting reports concerning the potential contribution of smoking and subsequent effects on sperm quality. Heavy cigarette smokers (>19 cigarettes/day), compared to lighter smokers (<9 cigarettes/day), had a significantly lower total sperm count [62]. Heavy smokers also had a significantly lower total sperm count than non-smokers. However, heavy smoking did not appear to affect sperm density, motility, or morphology. The combination of heavy coffee drinking and heavy cigarette smoking inhibited sperm motility and increased the proportion of dead spermatozoa. Also potentially relevant in this regard is the effect of in utero exposure to maternal tobacco smoke and decreased semen quality in male offspring. An investigation of the possible role of maternal smoking on the reproductive system in adult male offspring found no significant effects of early exposure to maternal smoking on conventional measures of semen quality (i.e., volume, density, motility, and morphology). However, current smoking was associated with a significant decrease in the percentage of sperm with normal morphology [45]. Without consideration of such biologically plausible potential confounders, the possible influence of exposure to compounds identified as environmental estrogens on declining sperm counts is difficult to assess.

CONCLUSIONS

The available evidence is too limited to judge with any confidence whether sperm counts have declined during the past 50 years. Based on both animal and human data with DES, it is biologically plausible that in utero exposures to exogenous estrogenic compounds are capable of reducing sperm production in adult males. However, the apparent existence of a maternal dose threshold for DES-induced effects on sperm count undermines the likelihood that environmental estrogens, which are substantially less potent, are capable of causing similar effects. The meta-analysis by Carlsen et al. [33] is not an adequate basis from which to conclude that sperm counts have declined worldwide. The possible sources of error in that study are numerous and include methodological changes over time, selection bias, geographical variability, and statistical issues. While some additional data appear to support the central thesis advanced by Carlsen et al. [33], other data, in particular the large study by Fisch et al. [47], appear to re-

fute the hypothesis that sperm quality has declined. This matter can be resolved only after all available data are comprehensively assessed, with adequate accounting of the potential sources of bias and error noted in this review. If the decline in sperm counts is real, careful evaluations of possible causes will include potential sources of estrogenic exposure, as well as other potential confounding factors (e.g., smoking, drugs, age). Finally, the documented lack of effects on sperm counts caused by elevated postnatal human body burdens of TCDD and PCBs suggests that sperm counts have not been adversely affected by exposure to these chemicals at typical environmental levels.

REFERENCES

1. W. R. Kelce, E. Monosson, M. P. Gamcsik. *Toxicol. Appl. Pharmacol.* **126**, 276 (1994).
2. National Wildlife Federation. *Hormone Copycats,* Great Lakes Natural Resource Center, Ann Arbor, MI (1994).
3. F. S. Greenspan and J. D. Baxter. *Basic and Clinical Endocrinology*, 4th ed., Appleton and Lange, Norwalk, CT (1994).
4. R. M. Berne and M. N. Levy. *Physiology*, 3rd ed., p. 1071, C. V. Mosby, St. Louis, MO (1993).
5. K. Verdeal and D. S. Ryan. *J. Food Protect.* **42**, 577 (1979).
6. J. A. McLachlan. "Rodent models for perinatal exposure to diethylstilbestrol and their relation to human disease in the male", in *Developmental Effects of Diethylstilbestrol in Pregnancy*, A. L. Herbst and H. A. Bern (Eds.), p. 148, Thieme-Stratton, New York (1981).
7. A. M. Soto, T.-M. Lin, H. Justicia, R. M. Silvia, C. Sonnenschein. "An 'in culture' bioassay to assess the estrogenicity of xenobiotics", in T. Colborn and C. Clement (Eds.), *Chemically Induced Alterations in Sexual Development: The Wildlife/Human Connection,* p. 295, Princeton Scientific Publishing, Princeton, NJ (1992).
8. A. M. Soto, K. L. Chung, C. Sonnenshein. *Environ. Health Perspect.* **102**, 380 (1994).
9. K. W. Gaido, L. S. Leonard, S. Lovell, J. C. Gould, D. Babai, C. J. Portier, D. P. McDonnell. *Toxicol. Appl. Pharmacol*. **143**, 205 (1997).
10. S. F. Arnold, D. M. Klotz, B. M. Collins, P. M. Vonier, L. J. Guillette, J. A. McLachlan. *Science* **272**, 1489 (1996).
11. K. Ramamoorthy, F. Wang, I.-C. Chen, S. Safe, J. D. Norris, D. P. McDonnell, K. W. Gaido, W. P. Bocchinfuso, K. S. Korach. *Science* **275**, 405 (1997a).
12. J. Ashby, P. A. Lefevre, J. Odum, C. A. Harris, E. J. Routledge, J. P. Sumpter. *Nature* **385**, 494 (1997).
13. K. Ramamoorthy, F. Wang, I.-C. Chen, J. D. Norris, D. P. McDonnell, L. S. Leonard, K. W. Gaido, W. P. Bocchinfuso, K. S. Korach, S. Safe. *Endocrinology* **138**, 1520 (1997b).
14. J. McLachlan. *Science* **277**, 462 (1997).
15. W. J. Dieckmann and M. E. Davis. *Am. J. Obstet. Gynecol.* **66**, 1062 (1953).
16. R. J. Stillman. *Am. J. Obstet. Gynecol.* **142**, 905 (1982).
17. R. M. Giusti, K. Iwamoto, E. E. Hatch. *Ann. Intern. Med.* **122**, 778 (1995).
18. A. L. Herbst, H. Ulfelden, D. C. Poskanzer. *N. Engl. J. Med.* **284**, 878 (1971).
19. M. Bibbo and W. B. Gill. *Obstet. Gynecol.* **49**, 1 (1977).
20. E. D. Whitehead and E. Leiter. *J. Urol.* **125**, 47 (1981).
21. B. E. Henderson, B. Benton, M. Cosgrove. *Pediatrics* **58**, 505 (1976).
22. A. L. Herbst, R. J. Kurman, R. E. Scully. *Obstet. Gynecol.* **40**, 287 (1972).
23. A. L. Herbst, D. C. Poskanzer, S. J. Robboy. *N. Engl. J. Med.* **292**, 334 (1975).
24. A. L. Herbst and M. M. Hubby. *Am. J. Obstet. Gynecol.* **141**, 1019 (1981).
25. F. J. Leary, L. J. Resseguie, L. T. Kurland, P. C. O'Brien, R. F. Emslander, K. L. Noller. *J. Am. Med. Assoc.* **252**, 2984 (1984).
26. O. C. Hadjimichael and J. W. Meigs. *J. Natl. Cancer Inst.* **73**, 831 (1984).

27. D. Labarthe, E. Adam, K. L. Noller, P. C. O'Brien, S. J. Robboy, B. C. Tilley, D. Townsend, A. B. Barnes, R. H. Kaufman, D. G. Decker, C. R. Fish, A. L. Herbst, J. Gundersen, L. T. Kurland. *Obstet. Gynecol.* **51**, 453 (1978).
28. O. P. Heinonen. *Cancer* **31**, 573 (1973).
29. W. B. Gill, G. F. B. Schumacher, M. Bibbo. *J. Urol.* **122**, 36 (1979).
30. A. J. Wilcox, D. D. Baird, C. R. Weinberg, P. P. Hornsby, A. L. Herbst. *N. Eng. J. Med*. **332**, 1411 (1995).
31. M. P. Vessey, D. V. Fairweather, B. Norman-Smith, J. Buckley. *Br. J. Obstet. Gynecol.* **90**, 1007 (1983).
32. Report to the Medical Research Council, "The use of hormones in the management of pregnancy in diabetics", *Lancet* **269**, 833 (1955).
33. E. Carlsen, A. Giwercman, N. Keiding, N. E. Skakkebaek. *Brit. Med. J.* **305**, 609 (1992).
34. R. M. Sharpe. *J. Endocrinol.* **136**, 357 (1993).
35. R. M. Sharpe and N. E. Skakkebaek. *Lancet* **341**, 1392 (1993).
36. J. McLachlan, R. Newbold, B. Bullock. *Science* **190**, 991 (1975).
37. E. Seaman, N. Bar-Chama, H. Fisch. *Mediguide Urol.* **7**, 1 (1994).
38. K. A. Ginsburg and D. R. Armat. *Fertil. Steril.* **53**, 882 (1990).
39. L. Johnson. *J. Androl.* **7**, 331 (1986).
40. R. J. Levine, R. M. Mathew, C. Brandon. *N. Engl. J. Med*. **323**, 12 (1990).
41. M. D. Whorton and C. R. Meyer. *Fertil. Steril.* **42**, 82 (1984).
42. D. Lerda. *Am. J. Ind. Med*. **22**, 567 (1992).
43. F. M. Wittmaack and S. S. Shapiro. *Wisconsin Med. J.* **91**, 477 (1992).
44. J. MacLeod and Y. Wang. *Fertil. Steril.* **31**, 103 (1979).
45. J. M. Ratcliffe and B. C. Gladen. *Reprod. Toxicol.* **6**, 297 (1992).
46. C. A. Paulsen, N. G. Berman, C. Wang. *Fertil. Steril.* **65**, 1015 (1996).
47. H. Fisch, J. Feldshuh, E. T. Goluboff, S. J. Broder, J. H. Olson, D. H. Barad. *Fertil. Steril.* **65**, 1009 (1996).
48. K. D. Smith and E. Steinberger. "What is digosperma?" in *The Testis in Normal and Infertile Men,* P. Troen and H. R. Nankin, (Eds.), p. 489, Raven Press, New York (1977).
49. H. Fisch and E. T. Goluboff. *Fertil. Steril.* **65**, 1044 (1996).
50. A. Brake and W. Krause. *Brit. Med. J.* **305**, 1495 (1992).
51. P. Bromwich, J. Cohen, I. Stewart, A. Walker. *Brit. Med. J.* **309**, 19 (1994).
52. G. W. Olsen, K. M. Bodner, J. M. Ramlow, C. E. Ross, L. I. Lipshultz. *Fertil. Steril.* **63**, 887 (1995).
53. T. A. Mably, D. L. Bjerke, R. W. Moore, A. Gendron-Fitzpatrick, R. E. Peterson. *Toxicol. Appl. Pharmacol.* **114**, 118 (1992).
54. D. L. Bjerke and R. E. Peterson. *Toxicol. Appl. Pharmacol.* **127**, 241 (1994).
55. United States Environmental Protection Agency (EPA), *Health Assessment Document for 2,3,7,8-Tetrachlorodibenzo*-p-*dioxin (TCDD) and Related Compounds*, Vols. I–III, EPA/600/BP-92/001c (1994).
56. C. Wilker, L. Johnson, S. Safe. *Fundam. Appl. Toxicol.* **141**, 68 (1996).
57. W. H. Wolfe, J. E. Michalek, J. C. Miner, A. J. Rahe. "An epidemiologic investigation of health effects in air force personnel following exposure to herbicides: Extract, reproductive outcomes, executive summary, introduction and conclusions", in *Air Force Health Study*, Epidemiology Research Division, Armstrong Laboratory, Brooks Air Force Base, TX (1992).
58. E. A. Emmett, M. Maroni, J. Jefferys, J. Schmith, B. K. Levin, A. Alvares. *Am. J. Ind. Med.* **14**, 47 (1988).
59. S. Dalterio, F. Badr, A. Bartke, D. Mayfield. *Science* **216**, 315 (1982).
60. S. Dalterio and A. Bartke. *Science* **205**, 1420 (1979).

61. B. S. Ahluwalia, S. U. Rajguru, G. H. Nolan. *J. Androl.* **6**, 386 (1985).
62. S. Osser, A. Beckman-Ramirez, P. Liedholm. *Acta. Obstet. Gynecol. Scand.* **71**, 215 (1992).