Research letter:

On influencing population means

Slama R, Siroux V,

Inserm U823, Team of Environmental Epidemiology, Grenoble, France.

Epidemiology, May;**23**(3):501-3, 2012

Findings on the existence or the lack of temporal trends in sperm concentration at the population level have been used to discuss the role of environmental factors (including endocrine disruptors) on male fertility. An assumption sometimes made [1] is that temporal trends in biological parameters such as sperm concentration will parallel temporal trends in their risk factors. As we illustrate below, this assumption is a too simple view, outside the probably rare situation where one prevalent environmental factor has a major impact on the biological parameter considered.

Let us consider the case of an environmental factor A to which no pregnant woman was exposed before year t_0 and to which 40% of pregnant women were exposed at a later time t_1; we assumed that in utero exposure to this factor causes an average decrease of 20% in sperm concentration in adulthood among male offspring. We also considered factors B and C, supposed to have a much stronger impact at the individual level (sperm concentration decrease by 85%) and whose prevalence in pregnant women rose by 10% (factor B) or 60% (factor C) between t_0 and t_1. We estimated the population mean of sperm concentration in adulthood among men born at period t_1, assuming that either factor existed alone; we assumed lack of selection bias, measurement or random error, and of temporal trends in the prevalence of any other factor. Using a simple simulation approach (detailed in eAppendix), we also estimated the change in mean concentration assuming that *several* factors simultaneously impacted on sperm concentration independently.

Compared to the unexposed cohort of men born at t_0, the impact of factor A in men from t_1 birth cohort corresponded to a decrease in sperm concentration by 8% (Table). Sperm concentration decrease was 9% for factor B, while it reached 51% in the case of the more prevalent factor C (Table). Finally, when exposure to 5 factors, each having the same individual impact as factor A (sperm concentration change by -20%) simultaneously increased, sperm concentration decreased by 34% at the population level; 8 such factors were required to cause a population decrease by 49% (Table).

This study shows that a single factor with a moderate but realistic influence on sperm concentration at the individual level (-20%, comparable to the reported effect of in utero

exposure to tobacco smoke [2]), whose prevalence has increased significantly over time (+40%) would cause a relatively small decrease in sperm concentration at the population level (-8%). This is because the population impact of a single factor will remain lower than its individual impact, except if prevalence increases from 0 to 100%. Several independent factors with moderate effects in the same direction and with rather high increase in exposure prevalence might entail substantial changes in sperm concentration.

Although some chemicals with a strong impact at the individual level have been identified in occupational settings,[3] the prevalence of exposure in the general population probably remained low. Therefore, to our knowledge, a factor with such high impact and prevalence than those hypothesized here for factor C has to date not been identified. If one looks for causes of a strong reduction in sperm concentration at the population level, a *combination* of several factors, each having a limited impact at the individual level and whose prevalences simultaneously strongly increased appears like a more realistic explanation. We assumed that factors acted independently, but of course synergy between factors could also exist [4] (see eAppendix for an illustration). In another setting, a simulation study has quantified the impact of public health interventions on smoking prevalence.[5]

Data on temporal trends in outcome alone (without individual information on exposures), although very relevant in a public health perspective,[6] correspond to a simple form of ecological studies and are, outside specific settings,[7] generally very limited to draw strong conclusions regarding the influence of environmental factors. Indeed, several factors may have opposed impacts at the individual level or opposed temporal trends. For these reasons, conclusions on the impact of endocrine disruptors and other families of environmental factors on male fertility (or the lack thereof) should not be drawn from studies of temporal trends in male fertility parameters alone.

In order to characterize the impact of exposures during the developmental window, mother-child cohorts with a biomarker-based assessment of exposure during pregnancy and long-term follow-up constitute a more relevant tool. Such studies are currently very rare.[2, 8]

REFERENCES

1. Bonde JP, Host Ramlau-Hansen C, Olsen J. Trends in Sperm Counts. The Saga Continues. *Epidemiology.* 2011;**22**:617-9.

2. Ramlau-Hansen CH, Thulstrup AM, Storgaard L, Toft G, Olsen J, Bonde JP. Is prenatal exposure to tobacco smoking a cause of poor semen quality? A follow-up study. *Am J Epidemiol.* 2007;**165**:1372-9.

3. Whorton D, Krauss RM, Marshall S, Milby TH. Infertility in male pesticide workers. *Lancet.* 1977;**2**:1259-61.

4. Kortenkamp A, Faust M, Scholze M, Backhaus T. Low-level exposure to multiple chemicals: reason for human health concerns? *Environ Health Perspect.* 2007;**115 Suppl 1**:106-14.

5. Ahern J, Hubbard A, Galea S. Estimating the effects of potential public health interventions on population disease burden: a step-by-step illustration of causal inference methods. *American journal of epidemiology.* 2009;**169**:1140-7.

6. Buehler JW. Surveillance. In: Rothman KJ, Greenland S, Lash TL, editors. *Modern Epidemiology*. 3rd Edition ed: Lippincott Williams & Wilkins; 2008. p. 459-80.

7. Clancy L, Goodman P, Sinclair H, Dockery DW. Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study. *Lancet.* 2002;**360**:1210-4.

8. Mocarelli P, Gerthoux PM, Needham LL, et al. Perinatal exposure to low doses of dioxin can permanently impair human semen quality. *Environmental health perspectives.* 2011;**119**:713-8.

TABLE

Table: Sperm concentration in adulthood among men born at periods t_0 and t_1, assuming the existence of an environmental factor whose prevalence increased between t_0 and t_1.

Exposure factor and birth cohort	Exposure prevalence [a]	Sperm concentration in male offspring in adulthood (millions/ml)		
		Non-exposed	Exposed	Whole population (Change [b])
Factor A [c]				
Period t_0	0%	100	N.A.	100x1 + 80x0 =**100**
Period t_1	40%	100	80 (-20%)	100x0.6 + 80x0.4=**92 (-8%)**
Factor B [c]				
Period t_0	0%	100	N.A.	100x1 + 15x0 =**100**
Period t_1	10%	100	15 (-85%)	100x0.9 + 15x0.1=**91.5 (-8.5%)**
Factor C [c]				
Period t_0	0%	100	N.A.	100x1 + 15x0 =**100**
Period t_1	60%	100	15 (-85%)	100x0.4 + 15x0.6=**49 (-51%)**
5 independent factors [d]				
Period t_0	0%	100		**100**
Period t_1	40%	100	(-20%/factor)	**67 (-34%)**
8 independent factors [d]				
Period t_0	0%	100		**100**
Period t_1	40%	100	(-20%/factor)	**52 (-49%)**

N.A.: not applicable.

In utero exposure to the hypothetical environmental factor is assumed to decrease sperm concentration in adulthood by 20% (factor A) or 85% (factors B or C) on average at the individual level.

[a] Frequency of exposure to one single factor among women pregnant the corresponding year.

[b] For a prevalence change by Δp, a biological parameter with an initial value of C_0 and an individual impact of x on a multiplicative scale (that is, the parameter is multiplied by x in exposed subjects), the relative change in the mean of the biological parameter in population t_1 is obtained as $(C_1-C_0)/C_0 = \Delta p(x-1)$. In the case of factor A, this is 0.4(0.2-1)=-0.08, a decrease by 8%.

[c] The population is assumed to be exposed *to factor A only* or *to factor B only* or *to factor C only*.

[d] Several deleterious factors similar to factor A are assumed to act in men born at t_1, each having a prevalence of 40% and entailing a 20% decrease in sperm concentration in adulthood in subjects exposed in utero, with no effect measure modification between these factors (that is, the probability of exposure to each factor at t_1 was 40% and independent from exposure to the other factors, and there was no modification of the effect measure of any factor on sperm concentration by any other factor)(see eAppendix for details).

eAppendix:

Several time-varying risk factors are needed to influence the population mean of a biologic parameter

Content:

I. Overview

In order to simulate the impact at the population level of an increase in prevalence over time of several environmental factors on sperm concentration, we used a simulation approach. The main steps, described below, are:

1) Simulation of a cohort of n=10,000 men born at t_0, all considered unexposed to the environmental factors of interest;
2) Simulation of a cohort of n=10,000 men born at t_1, assuming that each one has a probability p_{A1}, p_{A2}… , p_{Aq} to be exposed to independent factors A_1, … A_q, respectively;
3) Comparison of the mean sperm concentration between both birth cohorts to provide an estimate of the impact of factors (A_1, … A_q) considered altogether.

These steps have been repeated considering between q=2 and q=10 factors.
Finally, we also provide an illustration of a situation in which synergy between exposure factors is assumed.

II. Detail of the steps of the simulation study

II.1) Step 1

We draw the sperm concentration Y_{t0} of a hypothetical population P_0 of n=10,000 men, assuming a normal distribution with mean m_0=100 (expressed in million spermatozoa/ml) and standard deviation 50, replacing negative values by 0.

$$Y_{t_0} = \max\left[\mathcal{N}(100, 50); 0\right]$$

This corresponds to the sperm concentration in adulthood of men born at t_0.

II.2) Step 2

We draw the sperm concentration in adulthood Y_{t1} of another hypothetical population P_1 of n=10,000 men born at $t_1>t_0$; we first assume the same distribution as for population P_0. This starting value of sperm concentration for a subject i of P_1 is called $Y^i_{t_1(start)}$.

We then assume that each man of t_1 birth cohort is exposed to the group of factors (A1, … A_q) with probability (p_{A1}, … p_{Aq}). Factors A_1, … A_q are binary (one is either exposed or non exposed, or, equivalently, has an exposure above or below a given threshold value). In our example, for convenience all probabilities are assumed to have the same value p= p_{A1}= … =p_{Aq} , arbitrarily set to 0.4.

Let $\left(\mathrm{I}^i_{\mathrm{A}_1},\ldots \mathrm{I}^i_{\mathrm{A}_q}\right)$ be the indicator variables indicating if subject i is exposed to factors (A_1, … A_q), with $\mathrm{I}^i_{\mathrm{A}_p}$ = 1 if subject i is exposed to factor A_p and 0 otherwise. The probability of being exposed to a given factor is independent of whether or not subject i is exposed to any other factor (that is, variables $\left(\mathrm{I}^i_{\mathrm{A}_1},\ldots \mathrm{I}^i_{\mathrm{A}_q}\right)$ are independent one from another). In other words, if we consider, say, in utero exposure to tobacco smoke and alcohol consumption in adulthood, then we assume that the probability to drink alcohol is the same in men non-exposed and in men in utero exposed to tobacco smoke.

Each factor A_p is expected to entail a decrease by 20% on average in sperm concentration. This is achieved by multiplying sperm concentration by a variable with an expected value of 0.8 for subjects exposed to A_p. In order to allow for some individual variability in the sensitivity to A_p, we assume the effect of A_p to entail a decrease of 0.2 in average, with standard deviation 0.1.

That is, the effect of factor A_p is assumed to follow a law:

$$\mathrm{Effect}A_p \sim \mathcal{N}(.8, .1)$$

The sperm concentration of each subject of birth cohort t_1 is estimated by multiplying the starting value Y_{t1}(start) of each man by the values of the variables EffectA_p corresponding to the effects of each factor the subject is exposed to:

$$Y^i_{t_1} = Y^i_{t_1(start)} \times \left(EffectA_1\right)^{I^i_{A_1}} \times \ldots \times \left(EffectA_q\right)^{I^i_{A_q}}$$

II.3) Step 3

Finally, we compare mean sperm concentration between populations P_1 and P_0. The simultaneous effect of exposures A_1,… A_q at the population level is estimated as the difference between the mean values of sperm concentration of populations P_1 and P_0.

Steps 1 to 3 are repeated for q=2 to q=10 exposure factors.

III. Results

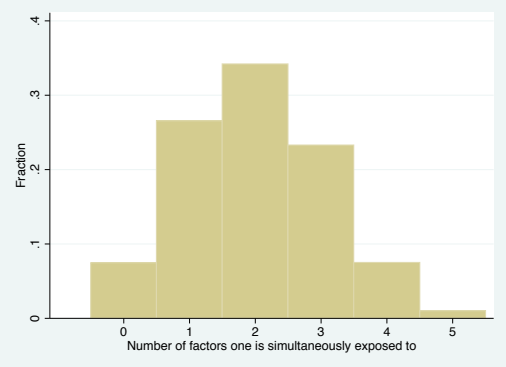
III.1) General case with independent environmental factors

The distribution of the number of factors each subject of birth cohort t_1 is simultaneously exposed to is given in Figure S1 for a maximum of 5 factors and a maximum of 10 factors. In the case when 10 factors exist, about 25% of subjects are exposed to 4 factors simultaneously, while 1.3% of subjects are exposed to 8 factors or more simultaneously.

The mean values of sperm concentration in population P_1 for q=1 to q=10 exposure factors is given Table S1. For 8 exposure factors, sperm concentration is on average decreased by 49% (mean, sperm concentration, 52 million/ml) compared to t_0 birth cohort (mean sperm concentration, 100 million/ml). For 10 exposure factors, sperm concentration is on average decreased by 57% (Figure S2).

Figure S1: Distribution of the number of factors one is simultaneously exposed to; **A)** Assuming that 5 different exposure factors exist; **B)** Assuming that 10 different exposure factors exist. Results are given for a specific simulation run and may slightly vary from run to run.

A) Up to 5 environmental factors



B) Up to 10 environmental factors

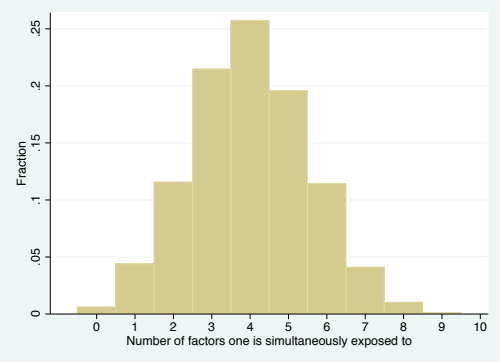
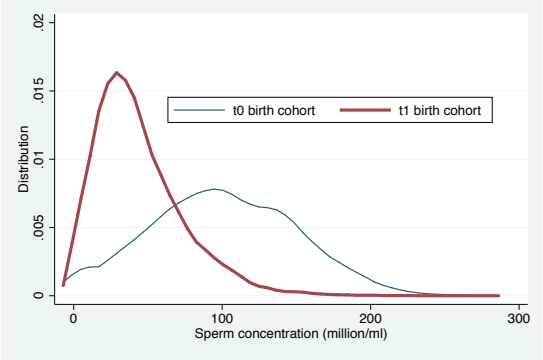


Table S1: Mean sperm concentration in adulthood in t_0 and t_1 birth cohorts, assuming that the birth cohort is exposed to 1 to 10 exposure factors with a negative individual impact on sperm concentration.

Number of exposure factors	Mean sperm concentration (sd), 10^6/ml		Estimated impact of exposures		
	Birth cohort t_0	Birth cohort t_1	Change in mean sperm concentration	Sperm concentration<10x10^6/ml, % (relative risk compared to t_0)	
0	*100 (49)*	*100 (49)*	*0*	*3.3*	*(1)*
1	100 (49)	92 (47)	-8%	3.9	(1.2)
2	100 (49)	86 (45)	-15%	4.1	(1.3)
3	100 (49)	79 (43)	-22%	4.3	(1.3)
4	100 (49)	72 (41)	-28%	4.7	(1.4)
5	100 (49)	67 (39)	-34%	5.1	(1.5)
6	100 (49)	61 (37)	-39%	5.4	(1.6)
7	100 (49)	56 (35)	-44%	5.9	(1.8)
8	100 (49)	52 (33)	-49%	6.5	(2.0)
9	100 (49)	48 (31)	-53%	7.3	(2.2)
10	100 (49)	44 (30)	-57%	8.4	(2.6)

Table S1 additionally gives the proportion of men with a sperm concentration below 10 million/ml according to the number of exposure factors considered, and the relative risk of having a sperm concentration below 10 million/ml in t_1 birth cohort, compared to the non-exposed birth cohort at t_0. Eight exposure factors induced a doubling in the proportion of men with a sperm concentration below 10 million/ml, compared to the unexposed cohort.

Figure S2: Distribution of sperm concentration in t_0 and in t_1 birth cohorts, assuming the existence of 10 independent environmental factors.



III.2) Further results in the presence of effect measure modification

The previous models assume that the probabilities of exposure to each factor are independent one from another and that the effect measure of a given factor on a multiplicative scale does not depend on the other factors one is exposed to.

Here, we relax the second hypothesis, starting from the situation with 5 exposure factors. Taken altogether and assuming independence, these factors entailed a decrease by 34% at the population level (Table S1). We now assume effect measure modification between factors A_4 and A_5, in that sperm concentration is further multiplied by 0.2 in men simultaneously exposed to factors A_4 and A_5. That is, the effect of A_4 (respectively, A_5) corresponds to a decrease by 20% in men not exposed to A_5 (respectively, A_4), and to a decrease by 87% (multiplication by 0.8x0.8x0.2 = 0.128) for men simultaneously exposed to these two factors.

In this first case with effect measure modification, the proportion of men of t_1 birth cohort simultaneously exposed to both factors was 16%; the simulation results predict a mean sperm concentration of 60 million/ml in men from t_1 birth cohort, a decrease by 40% compared to t_0 (this decrease was 34% without effect measure modification, Table S1).

If in addition we assume that in men simultaneously exposed to factors A_1, A_2 and A_3 sperm concentration is further decreased by 99% (or multiplied by 0.8^3x0.01=0.005 compared to unexposed men), then the predicted mean sperm concentration of men from t_1 birth cohort is 58 million/ml, a decrease by 42% compared to t_0. The proportion of men simultaneously exposed to all 3 factors was 5.9%.

IV. STATA code

```
*************************************************************************
***** Impact of an increased prevalence in environmental factors
*************************************************************************

*Last modified 23 Nov 2011.

version 11.2 /*This code: STATA v.11.2 format*/

clear
set seed 2011 /*Sets the root value of the random number generator*/

******************
*Step I: We generate sperm concentration of men born at period t0

set obs 10000 /*Generate 10,000 observations (without characteristics so far)*/
drawnorm conc, means(100) sds(50) /*Sperm concentration assumed to have normal distribution
with mean 100, sd 50.*/
replace conc=0 if conc<0 /*Negative values of sperm concentration are meaningless and set to
0*/

sum conc, d
hist conc
gen period=0 /*0=birth cohort t0 (observations 1 to 10,000 of dataset) */
label var period "Year of birth"

*Step II: We generate sperm concentration of men born at t1.
*First, we assume that the overall distribution is identical to that of men born at t0:
expand 2  /*We duplicate the dataset*/
replace period=1 if _n>10000   /*1=birth cohort t1 (Observations 10,001 to 20,000) */

drawnorm concb , means(100) sds(50)
replace concb=0 if concb<0
replace conc=concb if period==1

*We now successively assume that 1, 2... 10 environmental factors
* exist in at t1 and independently impact sperm concentration.
* Each factor implies a 20% decrease in sperm concentration (multiplication by .8 on average)
* and the prevalence of exposure is 40% at t1 for each factor (0% at t0).


**Option 1: No interaction; we assume that there is some variability in the effect of each
factor:

cap drop A* prob*
forvalues i=1(1)10 {
        gen A`i'=0 if period==1
        gen probA`i'=uniform() if period==1
        replace A`i'=1 if probA`i'>0.6 & probA`i'<1
        /*Ai variable is 1 for subjects exposed to factor Ai (40% of population) */
        drawnorm effectA`i', means(.8) sds(.1)
        /*This variable corresponds to the effect of factor Ai */
        replace conc=conc*effectA`i' if A`i'==1 & period==1
        /*Concentration if multiplied by effectAi for exposed subjects */
                *Step III: Comparison between both periods:
        dis "Number of environmental factors : " `i'
        ttest conc, by(period)
        }

gen nbfact=A1+A2+A3+A4+A5+A6+A7+A8+A9+A10
tab nbfact period, col
```

```
**Option 2: Interaction; we assume that there is some variability in the effect of each
factor:
replace conc=concb if period==1
cap drop A* prob*
forvalues i=1(1)5 {
       gen A`i'=0 if period==1
       gen probA`i'=uniform() if period==1
       replace A`i'=1 if probA`i'>0.6 & probA`i'<1
       /*Ai variable is 1 for subjects exposed to factor Ai (40% of population) */
       drawnorm effectA`i', means(.8) sds(.1)
       /*This variable effectAi corresponds to the effect of factor Ai */
       replace conc=conc*effectA`i' if A`i'==1 & period==1
       /*Concentration if multiplied by effectAi for exposed subjects */

       }

       sum effectA4 effectA5
*Additionally, we assume that sperm concentration is further reduced by 80% for subjects
simultaneously exposed to factors 4 and 5:
       replace conc=conc*.2 if A4==1 & A5==1 & period==1 /*Concentration if multiplied by
       effectAi for exposed subjects */
       dis "Number of environmental factors : 5 (with interaction between factors 4 and 5)"
       ttest conc, by(period)

*Additionally, we assume that sperm concentration is further reduced by 99% for subjects
simultaneously exposed to factors 1, 2 and 3:
       dis "Number of environmental factors : 5 (with interaction between factors (1, 2, 3)
       and (4, 5)"
       replace conc=conc*.01 if A1==1 & A2==1 & A3==1 & period==1
       ttest conc, by(period)
```