

with other accessible statistics, for the estimation of the contribution of the age trend and the regression contribution to the longitudinal change of a "high" subset which is being followed.

(E) Even if all of the analyses had been free of any defect the narrowness of the hypothesis being tested in the second paper—namely, that in the year when pollution decreases there is a decrease in the number of respiratory conditions which is greater in areas with the greater decrease in pollution, cannot justify the categorical conclusion that, "... the evidence suggests that the levels of pollution measured during the national study were not harmful to health." In the light of the findings in the first article¹ such a sweeping generalisation is unsupported.

Other workers have subsequently reported significant but small declines in a longitudinal study of pulmonary function in third, fourth, and fifth grade children with increases in 24 hour TSP and mean SO₂ in episodes in a range of approximately 0 to 275 µg/M³.⁵ A large scale French study in its cross sectional component found a statistical association between "upper respiratory tract" symptoms in children and sulphur dioxide over a range of daily means of 22 to 85 µg/M³ (specific method) or 13 to 127 µg/M³ (acidimetric method).⁶

Surely the importance of childhood respiratory disease and the high stakes and long term planning necessary in reducing pollution which may or may not be related to childhood pulmonary disease deserves a more straightforward and logical treatment of these complex problems. Since the data collected are of great value and most of the suggested approaches seem feasible, we urge that a more convincing approach be made to its analysis.

JOHN R GOLDSMITH
RAPHAEL TOEPLITZ

*Epidemiology and Health Evaluation Unit,
University Center for Health Sciences,
Ben-Gurion University of the Negev,
Beer Sheva 84 105, POB 653, Israel.*

Dr Melia and coauthors reply:

The letter of Professor Goldsmith and Mr Toeplitz raises several points that ought to have been clear from our original texts.^{1 2} Clearly all the results could not be presented in detail but those published do, we believe, give a fair representation of the results as a whole.

The first point concerning tables 5 and 6¹ arises because we did not make it clear enough that summer and autumn referred to the season of the interviews. The pollution values used were winter levels, as stated. Analyses with a smoke × SO₂ product term—to test for synergistic types of effect—showed nothing. We should have mentioned that.

The remaining points will be answered using the identifying letters of Goldsmith and Toeplitz.

(A) Weighting to combine several variables into one is inevitably controversial. Our position is that the simplest approach will generally be the most interpretable and a count (equal weights of one) is certainly easy to understand. Without general agreement about coding the severity of respiratory measures, we might have been severely criticised had we followed the suggestion of Goldsmith and Toeplitz.

(B) The table of individuals according to their initial counts of respiratory conditions in 1973 and the changes which occurred by 1974 may help to clarify this point. The children generated paired data which essentially required a form of McNemar's test for the analysis.⁷ This uses the elements falling on either side of the "no change" column to generate a proportion for analysis on the logistic scale. Admittedly this means combining several different types of cell—for example, 6-5, 4-3, and so on—but these are not as heterogeneous as Goldsmith and Toeplitz feared. Furthermore, as may be seen from the table, individual analysis of the groups defined by the number of initial symptoms would have been impractical because of the small numbers concerned.

(C) Goldsmith and Toeplitz have a point. Nevertheless, if such changes had been found they would have been evidence for the cause and effect

Number of children who had a decrease, no change, or an increase in number of respiratory conditions between 1973 and 1974, by the number of conditions reported in 1973

No of symptoms in 1973	Change in number of respiratory symptoms by 1974					Total
	Reduction of 2 or more	Reduction of 1	No change	Increase of 1	Increase of 2 or more	
0	0	0	516	63	13	592
1	0	66	64	12	5	147
2	19	14	21	11	3	68
3	11	5	9	2	1	28
4	9	2	1	2	1	15
5 or 6	1	3	3	0	0	7
Total	40	90	614	90	23	857

hypothesis. Nevertheless, the analyses were not restricted only to one year periods but, as indicated in the last paragraph of the results,² they also included data for a two year period. No evidence for an effect of pollution on health was found.

(D) Goldsmith and Toeplitz have misunderstood the application of work on regression to the mean. Unlike the study described by Chinn and Heller,⁴ in which changes in plasma cholesterol were analysed according to initial concentrations of plasma cholesterol, we did not select children for the analyses by initial number of respiratory conditions. The effect of age was considered in the longitudinal analysis and was discussed in the section "choice of data for analysis."²

(E) It should be clear that our conclusions are not as sweeping as Goldsmith and Toeplitz claim. We said that the evidence from our studies suggested that the pollution levels in the areas were below those necessary to harm health. Obviously our results are not proof of this and are subject to all the usual limitations of observational studies that Goldsmith and Toeplitz know as well as we do.

References

- ¹Melia RJW, Florey C duV, Swan AV. Respiratory illness in British schoolchildren and atmospheric smoke and sulphur dioxide 1973-7 I: cross-sectional findings. *J Epidemiol Community Health* 1981; **35**: 161-7.
- ²Melia RJW, Florey C duV, Chinn S. Respiratory illness in British schoolchildren and atmospheric smoke and sulphur dioxide 1973-7 II: longitudinal findings. *J Epidemiol Community Health* 1981; **35**: 168-73.
- ³Davis CE. The effect of regression to the mean in epidemiological and clinical studies. *Am J Epidemiol* 1976; **104**: 493-8.
- ⁴Chinn S, Heller RE. Some further results concerning regression to the mean. *Am J Epidemiol* 1981; **114**: 902-5.
- ⁵Dockery DW, et al. Change in pulmonary function in children associated with air pollution episodes. *J Air Pollut Control Assoc* 1982; **32**: 937-42.
- ⁶PAARC (Group Cooperatif). Pollution atmospherique et affections respiratoires chroniques ou a repetition: I Methods et sujets; and II Results et discussion. *Bull Eur Physiopathol Respir* 1982; **18**: 87-116.
- ⁷Breslow NE, Day NE. *Statistical methods in cancer research*. Vol I. *The analysis of case-control studies*. IARC scientific publications No 32. Lyon: International Agency for Research on Cancer, 1980.

Birth order as a quantitative expression of date of birth

SIR—The article by Berglin¹ exposing a persistently overlooked fallacy in the interpretation of alleged effects of birth order is to be welcomed as a major contribution on this difficult problem. The monograph² presenting the work in greater detail is essential for those subsequently working in this area,

and deals with many more pitfalls than it is possible to outline in a single paper.

It must be borne in mind, as Berglin himself points out,² that artefacts concerned in the estimation of birth order effects in adults are entirely different from those relevant to studies in the perinatal period, just as the methodology is necessarily different; the expression $(r-\frac{1}{2})/s$ is meaningful only in complete sibships. Indeed, in perinatal studies apparent parity effects must be assessed in the light of the fact that women with good reproductive histories tend to stop having pregnancies sooner than those with poor histories,³ yet this effect also invalidates longitudinal studies that control for the total number of pregnancies a woman has.⁴

The traits to which Berglin envisages his methods should be applied, such as eminence or alcoholism, have certain distinctive properties. They are not manifest at birth but appear at some (perhaps poorly defined) age of "onset" that varies between affected individuals. The traits are such that once such a label is attached to an individual it is not subsequently erased—in the terminology of stochastic processes, an "absorbing state."

Accordingly two effects may be expected:

(a) As pointed out by Berglin, in a study restricted to individuals born in a period in which the mean value of $(r-\frac{1}{2})/s$ is less than $\frac{1}{2}$ earlyborns with the trait will tend to outnumber lateborns, and vice versa.

(b) The relationship of predicted mean position (PMP) against year of birth is locally sinusoidal, with a similar, certainly non-trivial, derivative. So if a study population consists of individuals born during a period in which PMP varies across $\frac{1}{2}$, birth order will be confounded with age and hence with the degree of opportunity to have passed from the initial (negative) to the absorbing (positive) state.

To give a concrete example: the estimated number of heterozygotes for Huntington's chorea born each year in Glamorgan and Gwent did not show any clear trend during the period 1920-50.⁵ From Berglin's table 2 we may estimate a mean PMP of 0.532 for the period 1920-35, 0.464 for 1935-50, 0.498 for the whole 30 year period. A study based on births from 1920-35 would tend to show an excess of later born, while one for 1935-50 would tend to show an excess of first born. If we assume that the PMP values for Gothenburg apply to South Wales also, a study with ascertainment date 1980 based on births occurring during the whole period, considering a disease such as Huntington's chorea with a mean age at onset of around 45 years⁶ and a standard deviation of 12 years would be expected to give a weighted mean PMP, based on manifest cases, of 0.513. In this instance, effect (b), though smaller than effect (a), is not inconsiderable, and for conditions for which the