Journal of Epidemiology and Community Health, 1983, **37**, 78–81

Analysis of completed reproductive histories: a cautionary tale

JEAN GOLDING, [1] N R BUTLER,[1] AND R G NEWCOMBE[2]

From the Department of Child Health,[1] University of Bristol, Bristol, and the Department of Medical Statistics,[2] Welsh National School of Medicine, Cardiff, UK

SUMMARY Recent authors have suggested that cross-sectional studies of adverse outcome of pregnancy are misleading, and that the only valid method of analysis controls for eventual number of pregnancies. The present study shows, by simple examples, that such a method produces results that cannot be interpreted in the way claimed for them.

The fact that various adverse outcomes of pregnancy may be related to birth order has exercised epidemiologists, clinicians, and geneticists intermittently throughout the past 70 years.[1–5] On various occasions new methods of looking at the problem have emerged. From time to time it has been suggested that controlling for the total number of pregnancies that the woman eventually had would be more meaningful than any cross-sectional analysis.[6 7] This idea was revived by two publications using this method, one by Roman *et al*[8] analysing information obtained from women doctors and the other by Bakketeig and Hoffman[9] using linked birth records in Norway.

Man—or rather woman—is not an experimental animal. One cannot randomly mate 10 000 women on various occasions and document the outcomes of all the pregnancies that ensue. If it were possible to do this the problem of analysis and interpretation would be relatively simple. Reality is much more complex.

It is not our aim to produce an ideal method of analysis of birth order effects. On the contrary, we wish merely to point out that the so-called "longitudinal" method of analysis has a major flaw that makes interpretation of the data difficult. To show this we will generate two sets of reproductive histories that obey fixed rules. We will then apply the longitudinal method of analysis. If the method were valid those rules with which the data were generated should be shown.

In both data sets we make two fundamental assumptions:

(1) that women stop reproducing once their desired family size has been reached.

(2) that regardless of desired family size, there is restriction of fertility, unrelated to outcome of pregnancy. Such infertility might be due to either psychological, medical, or social factors.

Creation of the first set of reproductive histories

ASSUMPTIONS

(1) Suppose that the risk of fetal death is constant at 20% at each conception.

(2) Suppose that the desired family size consists of: one child in 5% of families, two children in 30% of families, three in 40%, four in 15%, and five children or more in 10% of the remainder.

(3) Suppose that the probability of ceasing to be fertile after any pregnancy is 15%.

Consider a population of 10 000 women expecting their first infant; on average 8000 of these pregnancies will result in live births and 2000 in fetal deaths. On average 400, 5% of the 8000 women, will be content with just one live child, and 1140, 15% of the remainder, will become infertile. Thus 1540 (400 + 1140) women will have a reproductive history of just one live birth, and 6460 will conceive again. Of the 2000 women who had a fetal death, an average of 300 (15%) will become infertile and 1700 will conceive again. Continuing in this way it is possible to work out the numbers of women who would have different reproductive histories (table 1).

RESULT OF CONTROLLING FOR NUMBER OF PREGNANCIES

The mainstay of the longitudinal method of analysis is to control for the total number of pregnancies and

Table 1 *Result of applying the first simple assumptions concerning risk of fetal death, desired family size, and secondary infertility to 10 000 hypothetical women*

Pregnancies	*Frequency*	*Pregnancies*	*Frequency*
One pregnancy only:		Four pregnancies only:	
D	300·0	DDDD	1·5
L	1540·0	DDDL	7·6
Two pregnancies only:		DDLD	5·6
DD	51·0	DLDD	5·6
DL	261·8	LDDD	5·6
LD	193·8	DDLL	62·5
LL	2162·4	DLDL	62·5
Thrre pregnancies only:		DLLD	15·3
DDD	8·7	LLDD	15·3
DDL	44·5	LDLD	15·3
DLD	32·9	LDDL	62·5
LDD	32·9	DLLL	275·1
DLL	367·6	LDLL	275·1
LDL	367·6	LLDL	275·1
LLD	90·2	LLLD	23·6
LLL	1618·4	LLLL	415·1
		Five pregnancies +:	1 404·9
		Total No of women	10 000·0

D = Fetal loss.
L = Live birth.

then to look for any birth order effect. Applying this method to the data of table 1 results in the rates of fetal loss shown in table 2 and depicted in fig 1. It may be seen from the right-hand column of table 2 that the overall rate of fetal loss is greatest for women who

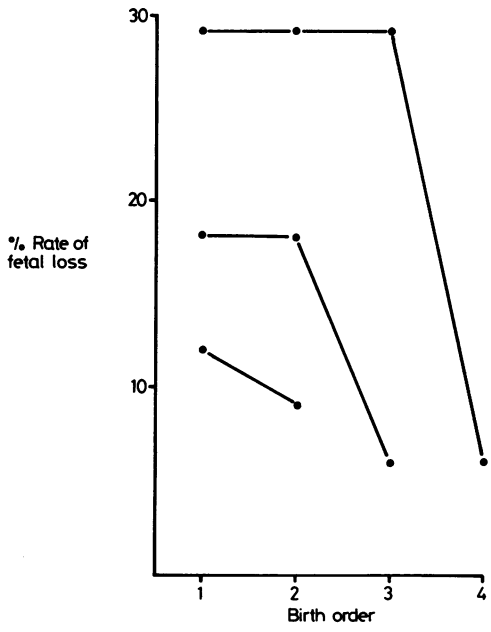

Fig 1 *Rate of fetal loss at each pregnancy order according to completed number of pregnancies; data generated using the assumption that the rate of loss is constant, at each pregnancy, for each woman but that women restrict their pregnancy history according to desired family size (first set of assumptions).*

Table 2 *Percentage fetal loss at each order of pregnancy according to total number of pregnancies, using first set of assumptions*

Final No of pregnancies	*Order of pregnancy*					
	1	*2*	*3*	*4*	≥5	All
1	16	—	—	—	—	16
2	12	9	—	—	—	10
3	18	18	6	—	—	14
4	29	29	29	6	—	23
Total*	20	20	20	20	20	20

*Including women of gravidity ≥5.

have the largest number of pregnancies and smallest for those who have just two.

Bakketeig and Hoffman[9] have interpreted such findings in a real study as meaning that women who have a large number of pregnancies have an inherently high risk of fetal death. But from the rules governing the generation of our data, the risk of fetal death was identical for all women in all pregnancies. The apparent contradiction is entirely a consequence of the fact that in our hypothetical situation women determine their reproductive history according to the number of children they desire. They do not have a fixed idea as to the number of times they wish to become pregnant.

Creation of the second data set

ASSUMPTIONS

(1) Suppose that the risk of fetal loss at birth order 1 is 20%, at birth order 2 it is 14%, at birth order 3 it is 20%, at birth order 4 it is 18%, and thereafter it is

20%, 22%, etc with successive pregnancies.

(2) Suppose that the same rules of infertility and desired family size apply as in the first data set.

If we generate obstetric histories as previously we obtain the distribution of completed pregnancies shown in table 3. Applying the method of analysis according to completed family size we obtain the pattern shown in table 4 and illustrated in fig 2. This is similar to the analyses of observed data (figs 3 and 4).

We would like to point out how misleading these pictures are. The authors of the analyses depicted in figs 3 and 4 claim that their data can be interpreted as showing that the risk to any particular woman falls with increasing birth order. Yet we have obtained similar results with data where, by definition, for each woman the risk with birth order was U-shaped.

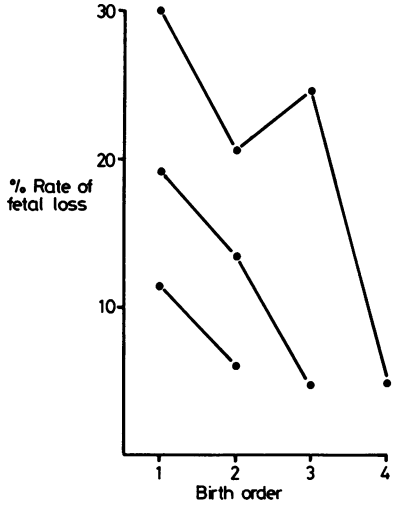

Fig 2 *Rate of fetal loss at each pregnancy according to completed number of pregnancies; data generated using the assumption that the rate of loss varies with pregnancy order, that no women are at greater risk than the others, but that women restrict their pregnancy history according to desired family size (second derived data set).*

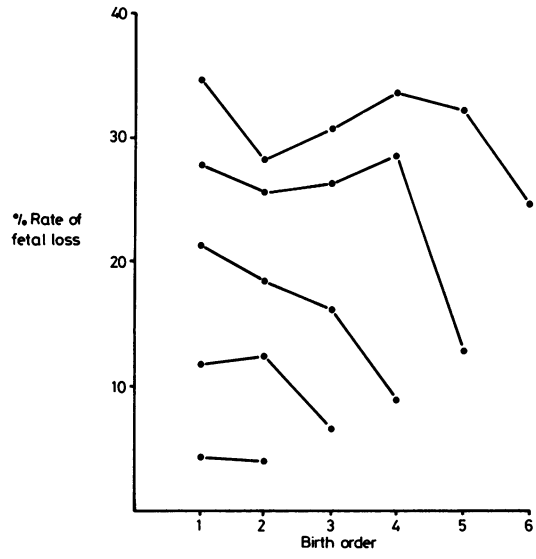

Fig 3 *Published rate of fetal loss by birth order in a population of women doctors,[8] analysed according to total number of pregnancies.*

Table 3 *Result of applying the second simple assumptions concerning risk of fetal death, desired family size, and secondary infertility to 10 000 hypothetical women*

Pregnancies	*Frequency*	*Pregnancies*	*Frequency*
One pregnancy only:		Four pregnancies only:	
D	300·0	DDDD	0·8
L	1540·0	DDDL	4·3
Two pregnancies only:		DDLD	3·7
DD	35·7	DLDD	4·3
DL	281·4	LDDD	2·8
LD	135·7	DDLL	47·1
LL	2324·5	DLDL	55·1
Three pregnancies only:		DLLD	15·6
DDD	4·9	LLDD	11·9
DDL	32·7	LDLD	10·1
DLD	28·3	LDDL	35·8
LDD	18·4	DLLL	318·3
DLL	415·0	LDLL	207·3
LDL	270·2	LLDL	242·5
LLD	77·6	LLLD	24·0
LLL	1826·8	LLLL	480·2
		Five pregnancies +:	1 245·0
		Total No of women	10 000·0

D = Fetal loss.
L = Live birth.

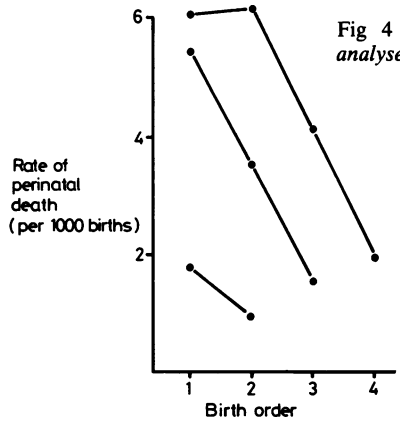

Fig 4 *Published rate of perinatal death in Norway,*[9] *analysed according to total number of births on file.*

Table 4 *Percentage of fetal loss at each order of pregnancy according to total number of pregnancies, using second set of assumptions*

Final No of pregnancies	*Order of pregnancy*					
	1	*2*	*3*	*4*	*5*	*All*
1	16·3	—	—	—	—	16·3
2	11·4	6·2	—	—	—	8·8
3	18·0	12·2	4·8	—	—	11·7
4	30·7	21·3	24·4	5·0	—	20·3
All women*	20·0	14·0	16·0	18·0	20·0	17·4

*Including women of gravidity ≥5.

Discussion

The purpose of this short paper has been to show that to control for the total number of pregnancies on the file is likely to produce a misleading pattern in any community where a measure of birth control is practised. The fact is that in Western civilisation women who start their obstetric histories with a disaster, or series of disasters, are likely to continue reproducing until they have produced one or more live births, whereas those who start with a successful outcome are likely to curtail their reproductive history as soon as they have reached their desired family size.

How then should one assess how the risk to the fetus varies with birth order? We would like to point out that the last row of each of tables 2 and 4 showed the true patterns that we had imposed on our generated data. It is unfashionable to praise the simple, but in this case the basic cross-sectional approach would have described the effects perfectly. We are not claiming that such an approach is the whole answer to this subject but that it holds the basis for further advance. A deeper analysis in preparation substantiates this. We hope we have shown clearly that the so-called longitudinal approach to this problem leads up a blind alley.

References

[1] Greenwood M, Yule GU. On the determination of size of family and of the distribution of character in order of birth from samples taken through members of the sibships. *Journal of the Royal Statistical Society* 1914; **77:** 179–97.
[2] Stevenson AC, Warnock HA. Observations on the results of pregnancies in women resident in Belfast: I. data relating to all pregnancies ending in 1957. *Ann Hum Genet* 1959; **23:** 382–94.
[3] Slater E. Birth order and maternal age of homosexuals. *Lancet* 1962; i: 69–70.
[4] Penrose LS. *The biology of mental defect.* 3rd ed. London: Sidgwick and Jackson Ltd, 1963.
[5] Barker DJP, Record RG. Birth order and the presence of disease. *Applied Statistics* 1969; **16:** 13–6.
[6] James WH. Stillbirth and birth order. *Ann Hum Genet* 1968; **32:** 151–62.
[7] James WH. Neonatal death and birth order. *Ann Hum Genet* 1970; **33:** 385–94.
[8] Roman EA, Doyle P, Beral V, Alberman E, Pharoah P. Fetal loss, gravidity and pregnancy order. *Early Hum Dev* 1978; **2:** 131–8.
[9] Bakketeig LS, Hoffman HJ. Perinatal mortality by birth order within cohorts based on sibship size. *Br Med J* 1979; iii: 693–6.