Brit. J. prev. soc. Med. (1972), **26**, 212-218

CYCLIC VARIATIONS IN DATE OF LAST MENSTRUAL PERIOD OF MOTHERS OF INFANTS WITH CONGENITAL MALFORMATIONS IN SOUTH WALES, 1964-66

C. J. ROBERTS, C. R. LOWE AND S. LLOYD

Department of Social and Occupational Medicine, Welsh National School of Medicine, Cardiff

In a review article on congenital malformations and season of birth, Bailar and Gurian (1965) stated that 'seasonal cycles in births of infants with certain malformations have been demonstrated beyond any reasonable doubt'. Recently, Wehrung and Hay (1970) have referred to the ambiguity of the term 'seasonal', and Hewitt *et al.* (1971) have pointed out that although a temporal pattern of behaviour may be demonstrated by an appropriate statistical test, the decision about whether the pattern is seasonal or not is, in the final analysis, subjective. The noun 'season' means a period of the year occurring naturally as a result of changes in the earth's position in relation to the sun, marked by changes in the length of day and night, and of climate. The adjective 'seasonal' is much less specific and means pertaining to, or associated with, one or more seasons. Epidemiologists have favoured the use of the term 'seasonal' to describe associations with temporal changes in an extremely wide range of variables from flora, fauna, insect vectors, recreational activities, and infectious diseases to critical dates in the tax calendar (MacMahon and Pugh, 1970). Within this range there would seem to be three broad categories of association (direct causal, indirect causal, and secondary (non-causal)):

(1) direct causal—in which the characteristic is directly influenced by season (e.g., the earth's position in relation to the sun, and the effect of this on temperature, rainfall, and length of day and night). Examples of illnesses which have direct temporal association are hypothermia in the elderly and heat stroke;

(2) indirect causal—in which the characteristic is related to season by way of an interaction with one of the direct characteristics (e.g., diet influenced by rainfall and temperature; clothing influenced by temperature; and outside recreational activities influenced by amount of daylight). Examples of illnesses which have indirect temporal associations are certain infectious diseases and skiing injuries;

(3) secondary—in which the characteristic is associated with season because both characteristic and season are independently related to a certain time of the year (e.g., the end of the tax year; biennial rate demands; school and public holidays). An example of this type of association is the increase in deaths from road traffic accidents at Christmas and other public holidays.

It is important to bear these three categories of association in mind since the temporal pattern of variables which are directly or indirectly associated may differ from that of variables which are secondarily associated. While each association could be described as showing 'cyclic' behaviour, the first two (direct and indirect causal) are likely to have a periodic function of one year's duration marked by a varying but smooth pattern which can be fitted to a simple harmonic curve (this we have called harmonic behaviour); while the third (secondary) is likely to have a periodic function of one year's duration marked by a significant change at a particular time of each year for a certain period of time (this we have called periodic non-harmonic behaviour). However, if (on inspection of each year's data) the significant change at a particular time of year (for a certain period of time) is observed for one year only (or appears but irregularly), we would describe the temporal behaviour as episodic.

An investigator concerned with the prevalence of congenital malformations who takes 'seasonal' to mean only the direct and indirect categories of association might therefore reasonably consider that fitting an harmonic curve is the appropriate test for a seasonal pattern, but one who includes secondary associations in his definition of seasonal may be justified in using the simple χ^2 test. If useful hypotheses are to be generated in this field it may be necessary not only to avoid the use of the confusing description 'seasonal' but also to particularize the observed cyclic characteristics of the data studied.

This problem did not arise in the early studies in this field (McKeown and Record, 1951) because the

changes in prevalence from month to month were so large and so regular that they could be demonstrated by relatively simple methods (plotting time of onset according to conventional time units, using a reference population to calculate rates where necessary) and were evident on visual inspection of the data.

However, rates from one part of the year to another became irregular and less obvious in the Birmingham data (from 1955) and in the Scottish data (from 1958), and subsequent studies in Birmingham, Scotland, and elsewhere have failed to demonstrate comparable findings (Leck and Record, 1966). Consequently, statistical methods have been proposed (e.g., Edwards 1961) which are capable of demonstrating harmonic fluctuations of a much lower order of magnitude than would appear significant in a simple χ^2 test. However, increased sensitivity is achieved at the expense of certain assumptions, which it may (or may not) be reasonable to make; e.g., Edwards' method assumes that the defect under investigation has a single high and a single low period of risk, and that these are approximately six months apart. Clearly, for conditions with more than one peak within a year (e.g., coarctation of the aorta—Miettinen, Reiner, and Nadas, 1970) this method would not be appropriate. Thus, when drawing conclusions about seasonality based on the results of these more sensitive tests, it is important that authors should state the assumptions they have made and their working definition of the term 'seasonal'. Unfortunately, this information is not available for most studies which have used such tests, with the result that their findings on seasonality are not strictly comparable.

In this paper we have attempted to make inferences on seasonality from the dates of the last menstrual period (LMP) for 24 categories of congenital malformations. In order to do this we have applied three statistical tests (simple χ^2; fitting an harmonic curve (Edwards, 1961); and the rank sum method (Hewitt *et al.*, 1971)) to each category of malformation, and we have then visually inspected the temporal behaviour of those categories of malformation which gave statistically significant results by any of the tests. In the discussion we have attempted to distinguish between those observations which identify the presence and nature of cyclic behaviour and the subsequent subjective considerations which may suggest that a particular temporal behaviour pattern is 'seasonal' in character.

Materials and Methods

Material

For three years beginning 1 January 1964, information was collected about all infants born to women resident in South Wales and about all the congenital defects identified in that birth population (Richards and Lowe, 1971). The survey area comprised two counties (Glamorgan and Monmouthshire) and the four county boroughs within their boundaries (Cardiff, Swansea, Merthyr Tydfil, and Newport). In the three years of the investigation, 92,982 infants (live and stillborn) were born to women resident in the area; of these infants, 90,921 were the outcome of singleton pregnancies and 2,061 of multiple pregnancies. In this paper we confine our attention to the singleton infants, and the reported analyses are based on the incidence of congenital defects as ascertained among them by the second anniversary of their births.

Methods of Statistical Analysis

(a) First a simple χ^2 test was applied to quarterly frequencies (i.e., quarters beginning in January, April, July, and October) with three degrees of freedom.

(b) Then the test which assumes a cyclic trend of a simple harmonic form was applied (Edwards, 1961). The analysis involves estimating two parameters, α (the amplitude) and ϕ (the position of the highest incidence). It also provides a test of significance of the trend using a χ^2 distribution with two degrees of freedom. Before using Edwards' formula the frequencies by month of LMP were adjusted for the different number of days in the month, and the resulting monthly distribution of each malformation was then further adjusted for seasonal variation in all births in the study area.

(c) Finally, we used the method which gives a rank to each month so that the month of the LMP with the highest incidence will have 12 and the lowest 1 (Hewitt *et al.*, 1971). In this method an attempt is made to define a distribution of the largest rank sum for any six-month segment by 5,000 Monte Carlo trials. We have used this distribution from Hewitt's paper to test our own data.

Results

Table I shows the significance levels for each category of malformation by each of the three statistical tests used. One category ('rest of CNS') was significant by the χ^2 test alone; one (hypospadias) by all three methods; one ('rest of limbs') by both Edwards' and the rank sum methods; two (septal defects and cleft palate alone) by Edwards' method alone; and two ('rest of alimentary' and talipes) by the rank sum method alone. The categories significant by Edwards' method (6, 11, 19, and 22) show a much greater level of significance than do those by the rank sum method (15, 18, 19, and 22).

TABLE I

STUDY OF POSSIBLE SEASONAL TRENDS IN DATES OF CONCEPTION (SOUTH WALES DATA 1964-66); COMPARISON OF RESULTS BY SIMPLE χ^2, EDWARDS' METHOD, AND RANK SUM METHOD

Category of Malformation	No. of Cases	χ^2	Edwards' Method		Rank Sum Method
		Value of χ^2 (df=3)	Value of χ^2 (df=2)	Month of Highest Incidence	Rank Sum
1. Anencephalus	253	1·640	0·243		43
2. Spina bifida	311	7·290	4·739		48
3. Hydrocephalus	75	1·686	2·191		48
4. Rest of CNS	20	8·273*	5·698		52
5. All of CNS	659	5·439	2·615		50
6. Septal defect alone	65	6·805	14·184‡	September	50
7. Patent ductus alone	32	1·566	0·235		47
8. Rest of heart	252	4·107	3·285		48
9. All heart	349	4·951	1·785		49
10. Cleft lip alone	46	6·028	0·741		43
11. Cleft palate alone	66	6·863	7·537*	March	52
12. Cleft lip with palate	56	4·372	5·738		52
13. Atresia	78	3·197	1·963		50
14. Pyloric stenosis	181	4·423	1·125		49
15. Rest of alimentary	84	1·550	2·522		55*
16. All alimentary	511	2·640	1·654		48
17. Congenital dislocation of hip	74	1·516	0·838		49
18. Talipes	282	1·946	3·948		55*
19. Rest of limbs and skeleton	335	7·470	11·455†	June	57*
20. All limbs and skeleton	691	0·822	0·626		47
21. Mongolism	79	1·287	2·133		49
22. Hypospadias	102	13·848†	14·125‡	November	55*
23. Other urogenital	115	1·580	2·305		54
24. Skin	339	2·106	1·289		49

Significance levels: * $P < 0·05$
† $P < 0·01$
‡ $P < 0·001$

The Figure shows the ratio of observed to expected monthly numbers for each defect which was significant by any of the tests applied. Visual presentation of the temporal behaviour of the data helps to explain much of the lack of concordance between individual tests. The display for 'rest of CNS' (Figure A) does not follow the pattern of a simple harmonic curve, nor is there an indication of a run of six month highs followed by six month lows. Unfortunately, the small total number of defects in this category (n = 20) precludes breakdown by individual year, and we are thus unable to determine whether the observed pattern is periodic (non-harmonic) or episodic (i.e., occurring in one year only, or irregularly).

Septal defects and cleft palate alone (Figure B and C) show a simple harmonic pattern on visual display and were significant by Edwards' method. But because of meandering (sudden disturbances away from the trend) they failed to demonstrate the harmonic pattern demanded by the rank sum method (namely a run of six month highs followed by a run of six month lows). In contrast, talipes and 'rest of alimentary defects' (Figure D and E) showed an harmonic pattern on visual display, were not significant by Edwards' test, but both had a run of highs followed by a run of lows (without meandering) and were therefore significant by the rank sum method.

Two defects, 'rest of limbs' and hypospadias (Figure F and G), were significant by both methods and on visual display they show a simple harmonic pattern which satisfies both tests—a high peak separated from a low peak by six months with one cycle in the year, with very little sudden disturbance of high or low runs. We can conclude with reasonable confidence that the temporal behaviour of these defects is truly harmonic. The problem of interpreting results which are significant by one test only is examined in the discussion.

DISCUSSION

Before proceeding to discuss the inferences that can be drawn from any particular set of temporal characteristics in respect of the likely nature of a seasonal association (i.e., direct, indirect or secondary) it is important first to consider the data to which the statistical tests have been applied. In our study, three of the seven categories of malformation which show a significant cyclic trend by at least one test consist of groups prefixed by 'rest of'. Because of their heterogeneity none of the categories would be expected to have a common aetiology. When three tests are applied to each of 24 categories it is to be expected that three should show significance at the 5% level. This may explain the findings for 'rest of CNS' and 'rest of alimentary'. With regard to 'rest of limbs and skeleton', however, we suspect that the variation shown may

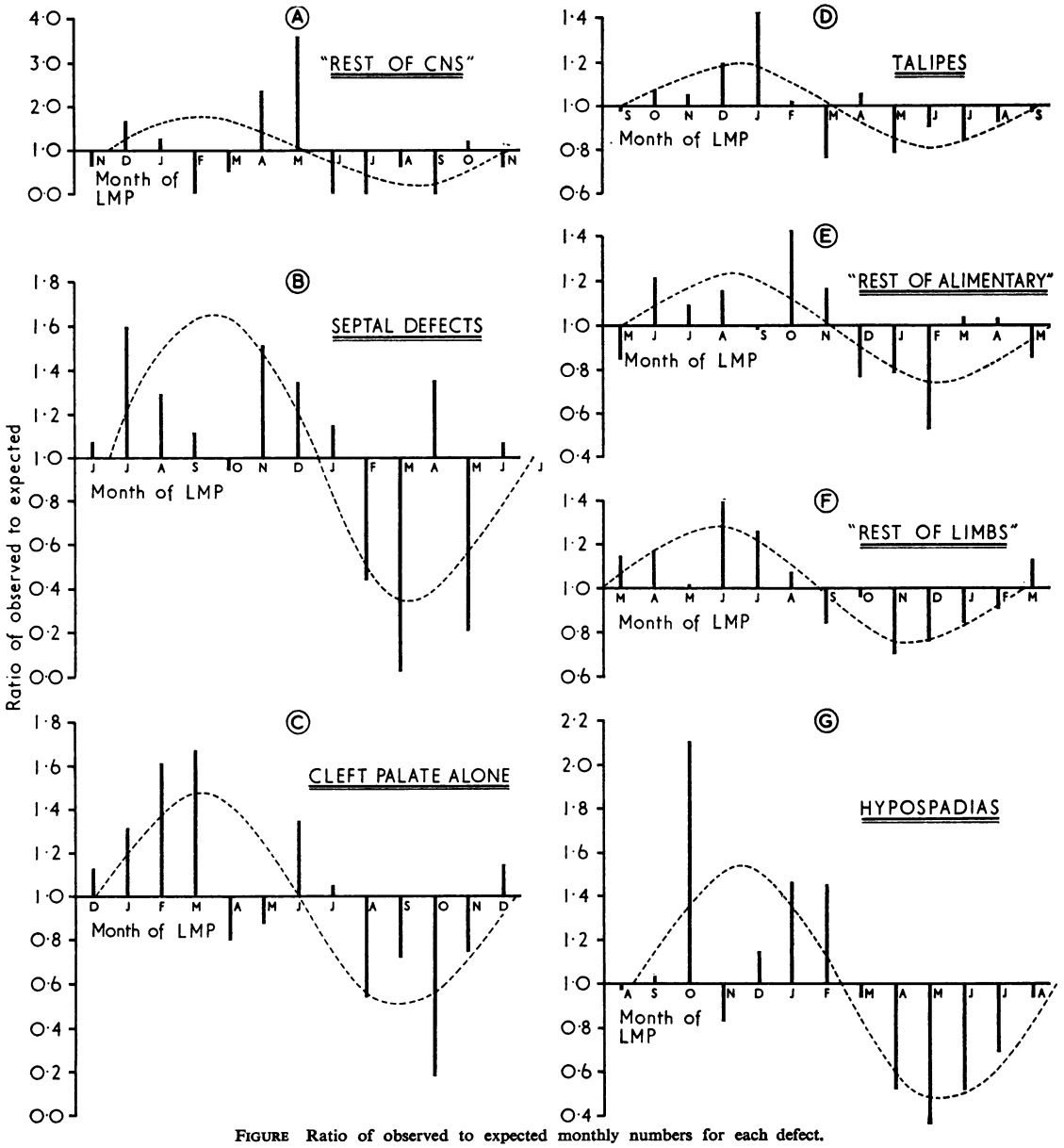

FIGURE Ratio of observed to expected monthly numbers for each defect.

be a real one due possibly to the influence of a subgroup within this category (probably polydactyly, syndactyly, and reduction deformities) which has a strong cyclic behaviour. This is being investigated.

Having drawn attention to those heterogeneous categories where the results of statistical tests should be accepted with caution, we wish to discuss the possible significance of the observed differences between the findings of individual statistical tests. Our analysis suggests that the three tests used are not comparable. Consequently, if (as has been the tendency in the past) seasonality has been inferred from a significant result by any one of them, then different workers must hold different views about what 'seasonal' means. This makes comparison between studies impossible.

A χ^2 test on our data (Table I) showed both hypospadias and 'rest of CNS' to be significant at the 5% level (in 24 categories it is to be expected

that one would be significant at the 5% level by chance alone). However, 'rest of CNS' was not significant by the remaining two tests, unlike hypospadias which was significant at 0·1% by Edwards' method and at 5% by the rank sum method.

A simple χ^2 test can show differences in time if the subgroup numbers are large enough (and thus can be influenced by the way in which the subgroups are chosen), but it does not take into account any pattern or trend within the data studied; a series of monthly frequencies could be scattered above and below the mean, and if the differences were large enough the test would show statistical significance. Clearly, an important issue with respect to interpretation of significant findings from a simple χ^2 analysis on results which are the amalgamation of several years' data is to determine (by re-examination of the results for each year of the study) whether the observed change occurs at the same time each year (periodic non-harmonic) or in one year only (episodic). The former temporal pattern is likely to generate a different set of hypotheses from the latter.

Edwards' test showed four malformations to be significant at the 5% level, but it is questionable whether it is reasonable to fit a harmonic curve to two of these (septal defects and cleft palate alone) in view of the meandering of monthly frequencies shown in the visual display (Figure B and C).

Edwards (1961) acknowledged that his method does allow some meandering in the detection of simple harmonic patterns; he stated 'it may also conceivably happen that there will be an increase or decrease in one month due to a sudden disturbance, which is not repeated in other years. It is not possible to eliminate the effect of such disturbances completely, but the use of a square root transformation may help to reduce them'. An example of such a disturbance can be found from his data on monthly number of anencephalic first births in Birmingham (Edwards, 1961):

Numbers ..	10	19	18	15	11	13	7	10	13	23	15	22
Ranks ..	2	10	8*	7	4	5	1	3	6	12	9*	11

* The rank order is reversed because total births for March are much greater than for November.

The relative amount of meandering in any one year is an important criterion for determining the suitability of Edwards' test, and it may well be that before any 'significant' results are considered further the data to which they have been applied should be visually displayed, and if meandering is found in more than one or two months of the year the test may not have been appropriate. This suggestion is supported by the application of the rank sum method to Edwards' Birmingham data for anencephalic first births (the rank sum is $12 + 9 + 11 + 2 + 10 + 8 = 52$ and gives a result which is not significant), yet Edwards' method yields a 5% significance (a similar situation exists for his later anencephalic births). It is interesting to note that in our own data 'rest of limbs' and hypospadias (which displayed much less meandering than did the above defects and, therefore, closely satisfied the criteria proposed by Edwards) were also significant by the rank sum method at the 5% level. This suggests that the rank sum method could possibly be used as a check on the appropriateness of Edwards' method, thus avoiding recourse to a visual display, and vice versa.

When defects show a discordance between results by Edwards' and the rank sum methods the findings are very difficult to interpret. Table II shows a comparison of the two methods at significance levels of 5 and 10% as applied to the malformations in Table I. Hewitt *et al.* (1971) stated, on the basis of the Monte Carlo analysis, that the assumption of χ^2 as a null distribution is not justified when the sample number is small. Their findings show that a marked excess of significant results appears when sample sizes fall below 72. We have, therefore, excluded malformation categories with less than 70 cases from the comparison. Of the 18 categories of malformation to which the two tests were applied, 12 were below the 10% level of significance by both

TABLE II

COMPARISON OF EDWARDS' TEST AND RANK SUM METHOD AT SIGNIFICANCE LEVELS OF 5% AND 10% APPLIED TO ALL MALFORMATIONS IN TABLE I, BUT EXCLUDING THOSE CATEGORIES WITH LESS THAN 70 CASES

		Significance Levels by Edwards' Test		
		$P < 0·05$	$0·05 \leqslant P < 0·1$	$P \geqslant 0·1$
Significance levels by rank sum method	$P < 0·05$	2(19)*, 2(22)		2(15), 2(18)
	$0·05 \leqslant P < 0·1$			1(23)
	$P \geqslant 0·1$		1(2)	12

*Malformation category number in parentheses (refer to Table I).

tests. One (spina bifida) was significant at the 10% level by Edwards' test but not by the rank sum method; one ('other urogenital') was significant at the 10% level by the rank sum method but not by Edwards' test; two ('rest of alimentary' and talipes) were not significant by Edwards' test but were significant at the 5% level by the rank sum method, and two (hypospadias and 'rest of limbs') were significant at the 5% level by both tests.

At present we feel that both methods should be used, and that only defects which give significant results by both should be regarded as having definite harmonic behaviour. In our data, therefore, we regard only hypospadias and 'rest of limbs' as showing harmonic behaviour, and we reserve judgement on the others (talipes, 'rest of alimentary', septal defects, and cleft palate) until more evidence is available from other studies. Support for our claim that hypospadias has an harmonic pattern comes from a recent study by Wehrung and Hay (1970), based on data from the National Cleft Lip and Palate Intelligence Service, U.S.A.

In summary, the simple χ^2 test detects changes in frequency in relation to the amplitude of the changes rather than to any temporal pattern. Edwards' method can detect a simple harmonic pattern (this assumes that the peak and the trough of an annual cycle are separated by approximately six months); it may also detect recurrent outbreaks (Wehrung and Hay, 1970) but it gives higher levels of significance than it should when the sample size is small (Hewitt *et al.*, 1971); and it requires assumptions on meandering of the data. The rank sum method also detects simple harmonic patterns but is more critical to meandering, can distinguish between harmonic and periodic (non-harmonic) variation, and detects much smaller sizes of differences between highs and lows (the value, if any, of the latter capability remains to be assessed).

As Hewitt *et al.* pointed out, statistical tests merely detect significant variation between months, and it remains with the investigator to judge whether the variation is seasonal in character. It has been suggested that this problem can be ameliorated when certain situations (e.g., if a prior hypothesis is available) allow the use of more efficient tests. However, it is unlikely that an efficient test does any more than bring to light significant variations which were not evident by less efficient tests, and it will not help the investigator in his subjective judgement about whether the variation is 'seasonal' in character. In the situation most frequently encountered in the epidemiology of congenital malformations, there is usually no prior hypothesis, and consequently a battery of tests for cyclic behaviour is applied (rather like a screening test) to all categories of malformation in a particular survey. Subsequently, the term 'seasonal' is used as a blanket description for all defects which show significant temporal variation (irrespective of the nature of the test used). An example of this approach will be found in a review of congenital malformations and season of birth (Bailar and Gurian, 1965).

We suggest that the term 'seasonal' should be avoided in epidemiological descriptions because it does not particularize the temporal characteristics of the data and so is less likely to produce specific hypotheses than a description which does. For example, from our own study hypospadias might have been described simply as showing a seasonal increase among winter conceptions; much more informatively, it can be described as showing a marked harmonic pattern (detected by two independent tests and confirmed on visual display) which clearly shows both a winter peak and a summer trough with no sudden disturbances away from that trend. A description such as this may well lead to a more specific aetiological hypothesis than would the mere description 'seasonal'. For example, it could be argued that such close adherence to an harmonic pattern (which was also present for each year of the study—1964, 1965, and 1966) suggests a direct or indirect causal (e.g., hours of daylight or temperature) rather than a secondary association with season. Such reasoning is consonant with studies of the defect in experimental animals; an association has been demonstrated between hours of daylight and pituitary function in birds (Hammond, 1954) and ferrets (Clarke, McKeown, and Zuckerman, 1939), and between pituitary function and hypospadias in the rabbit fetus (Jost, 1965).

Summary

The temporal behaviour of dates of the last menstrual period of mothers of infants comprising 24 categories of congenital malformations (derived from 90,921 singleton births in South Wales, 1964-66) are examined. Three statistical tests are applied to each category of malformation (simple χ^2; fitting an harmonic curve (Edwards, 1961); and the rank sum method (Hewitt, Milner, Csima, and Pakula, 1971). The temporal behaviour of those categories of malformation which gave statistically significant results by any of the tests is then examined visually. Only the categories of hypospadias and 'defects of limbs and skeleton (other than congenital dislocation of the hip and talipes') showed harmonic behaviour (we reserve judgement on talipes, 'rest of alimentary conditions', septal defects, and 'cleft palate alone').

In the discussion we have attempted to distinguish between those observations which identify the presence and the nature of cyclic temporal behaviour among congenital malformations, and the subsequent subjective considerations which may suggest that a particular temporal pattern is 'seasonal' in character. Our findings lead us to conclude that (1) the term 'seasonal' should be avoided in epidemiological descriptions because it does not particularize the temporal characteristics of the data and so is less likely to produce a specific hypothesis than a description which does; and (2) the term 'cyclic' should be prefixed by a description of its function, i.e., harmonic or periodic (non-harmonic), since this is also likely to assist in the generation of a more specific hypothesis.

We acknowledge with gratitude the generous grant from the Association for the Aid of Crippled Children, New York, which made this investigation possible. It is with pleasure that we also acknowledge our indebtedness to the Medical Officers of Health of Cardiff, Merthyr Tydfil, Newport, Swansea, Glamorgan and Monmouthshire for their interest and co-operation and to the many consultant obstetricians, paediatricians, pathologists, dermatologists, and orthopaedic, plastic, and ophthalmic surgeons who have notified and continue to notify to us the names of infants in our birth population who have come to their attention with congenital defects. We are also grateful to Dr. J. Howlett and Mrs. J. Lay, of the Atlas Computer Laboratory, Chilton, Berks, for the help they and their staff have always given us.

REFERENCES

BAILAR, J. C., and GURIAN, J. (1965). Congenital malformations and season birth. A brief review. *Eugen. Quart.*, **12**, 146.

CLARKE, W. E. LE GROS, MCKEOWN, T., and ZUCKERMAN, S. (1939). Visual pathways concerned in gonadal stimulation in ferrets. *Proc. roy. Soc. B*, **126**, 449.

EDWARDS, J. H. (1961). Seasonal incidence of congenital disease in Birmingham. *Ann. hum. Genet.*, **25**, 89.

HAMMOND, J. (1954). Light regulation of hormone secretion. In, *Vitamins and Hormones:* Vol. 12, p. 157. Edited by R. S. Harris, G. F. Marrian, and K. V. Thimann. Academic Press, New York.

HEWITT, D., MILNER, J., CSIMA, A., and PAKULA, A. (1971). On Edwards' criterion of seasonality and a non-parametric alternative. *Brit. J. prev. soc. Med.*, **25**, 174.

JOST, A. (1965). Gonadal hormones in the sex differentiation of the mammalian foetus. In: *Organogenesis*, edited by R. L. Dettaan and H. Ursprung. Holt, Rinehart, and Winston, New York.

LECK, I., and RECORD, R. G. (1966). Seasonal incidence of anencephalus. *Brit. J. prev. soc. Med.*, **20**. 67.

MACMAHON, B., and PUGH, T. F. (1970). In: *Epidemiology, Principles and Methods.* Little Brown, Boston.

MCKEOWN, T., and RECORD, R. G. (1951). Seasonal incidence of congenital malformations of the central nervous system. *Lancet*, **1**, 192.

MIETTINEN, O. S., REINER, M. L., and NADAS, A. S. (1970). Seasonal incidence of coarctation of the aorta. *Brit. Heart J.*, **32**, 103.

RICHARDS, I. D. G., and LOWE, C. R. (1971). Incidence of congenital defects in South Wales, 1964-66. *Brit. prev. soc. Med.*, **25**, 59.

WEHRUNG, D. A., and HAY, S. (1970). A study of seasonal incidence of congenital malformations in the United States. *Brit. J. prev. soc. Med.*, **24**, 24.