ON THE NATURE OF THE DETERMINANTS OF CONGENITAL PYLORIC STENOSIS

BY

GEORGE KNOX*

From the Department of Social Medicine, University of Birmingham

A comprehensive hypothesis of the aetiology of congenital pyloric stenosis must account for a number of striking and well established features. They include a differential sex incidence, the disease being four to five times more frequent in boys than in girls; a falling incidence with increasing birth rank; a familial concentration of the disease in the parents, siblings, and offspring of affected individuals (McKeown, MacMahon, and Record, 1951a,b; McKeown and MacMahon, 1955).

There is another feature of the disease which has lately come to notice (Table I). It has been shown that there are more females than males among the sibs of parents of affected children and that, although this disturbance is present among the sibs of both parents of affected girls, it is found only among the mothers' sibs when the affected children are boys (Knox, 1958a). The finding therefore is limited to sibships from which the affected children drew X-chromosomes. This recalls at once the striking sex ratio of the affected children themselves and suggests that a sex-linked genetic factor may be responsible for both features.

If the situation were indeed so simple that it could be explained in terms of an X-borne recessive gene with equal penetrance of the two appropriate genotypes, then the ratio of females affected to males affected would provide a direct measure of the gene frequency (g). This is because the frequency of the XY genotype in males is g and that of the XX genotype in females is g^2. Since congenital pyloric stenosis occurs in males and females in the ratio 5:1, the frequency of the hypothetical determining gene would be about 0.2. The immediate implication is that one-fifth of boys and one-twenty-fifth of girls are genetically susceptible to the disease. This amounts to 120 per 1,000 children of both sexes, and since the observed incidence of the disease is

about 3 per 1,000 live births the penetrance factor of the appropriate genotypes is about 1/40.

To this extent the hypothesis is entirely consistent with evidence of the importance of non-genetic factors in the aetiology of the disease (McKeown, MacMahon, and Record, 1952). Even if the accuracy of the gene-frequency estimate were upset by unequal penetrance in the two sexes, the incidence of the disease and the evidence of important environmental factors together confirm the general terms of the hypothesis to be tested.

Hypotheses of this nature usually require an apology, for they contain enough variables to fit many different assemblies of data, but in the present case the suggestion is based upon a most remarkable pattern of events whose explanation in other terms has so far proved elusive. The further purpose of this paper is to test the hypothesis by examining its corollaries in relation to the observed facts and to reassess its plausibility in terms of what is found. Many of the corollaries are deduced algebraically and as the deductions are rather complex and the results of more general interest they are the subject of another paper (Knox, 1958b).

FREQUENCY OF THE DISEASE IN RELATIVES

(*i*) *Cousins.*—The expected incidence of disease among cousins of children with a disease determined in this manner supplies a basis for a test. With a gene of frequency 0.2 we would not expect a greatly increased total incidence: precisely, $4.4/1,000$ cousins against $3/1,000$ of the population [Appendix]. Large numbers would be necessary to detect a difference of this degree and it is more relevant to the present purpose that we can expect the proportions of affected maternal and paternal cousins to be, respectively, 0.622 and 0.378.

Data from Birmingham (McKeown and others, 1951b) showed eleven affected among 3,980 cousins of

---

propositi. As anticipated, this fails to distinguish between the population incidence and that expected under the terms of the hypothesis, but eight of the eleven cousins were maternal cousins. I have found seven other papers in which the requisite details of the relationships between affected cousin pairs were given (Cockayne, 1934, 1938; Cockayne and Penrose, 1943; Carter and Savage, 1951; Fenwick, 1953; Carter and Powell, 1954; Nielson, 1954). By counting each affected child in turn as propositus, these authors' papers yield thirty maternal cousins and eighteen paternal ones. Addition to the Birmingham figures makes 41 maternal and 26 paternal compared with expected figures of 41·7 and 25·3 under the terms of the hypothesis. Although the ratio 41/26 does not differ significantly from unity ($\chi^2 = 3·36$; $0·05 < P < 0·1$), the theoretical consequences of a sex-linked gene with frequency 0·2 could scarcely correspond more closely with the facts.

(ii) *Parents*.—The expected proportion of mothers with genotype $\dot{X}\dot{X}$ is g and the proportion of $\dot{X}Y$ fathers is $2g/(1+g)$ (Knox, 1958b). With a gene of frequency 0·2, we expect a male/female ratio among affected parents of affected children equivalent to 5:3 instead of the population ratio of 5:1. McKeown and MacMahon (1955) reported twelve new families in which the disease occurred in two generations, sometimes with more than one child affected. If each affected child is counted once, these data give nine affected mothers and six fathers. Data supplied by Carter and Powell (1954), treated in a similar manner, give nine mothers and seven fathers. Nielson (1954) reported the families of nine affected parents in which one or more children (eleven altogether) were affected: the necessary information for counting twice the two parents with two affected children was not supplied, but there were three affected mothers and six affected fathers. Nielson also reported a further two affected fathers by tracing his families in the opposite direction. Fenwick (1953) reported a family with a father and two sons affected and a similar family has been reported anonymously (Anon., 1958). Altogether, in these five papers, 46 affected parent-child pairs could be identified and in 25 the affected parent was the father. Under the terms of the hypothesis the expected number is 28·7 and in terms of the population sex ratio it is 38·3. The difference between the latter expectation and the observations is statistically significant, but the facts are quite compatible with the hypothesis of a sex-linked determining gene with frequency 0·2.

However, the available data on parent-child pairs supply the first evidence against the plausibility of an explanation in terms of sex-linkage. McKeown and MacMahon (1955) supply six father-child pairs in which the sex of the child was reported, Carter and Powell (1954) supply seven, Fenwick (1953) two, and the anonymous author (1958) another two. Fourteen of the seventeen children were boys and only three were girls. This differs significantly from the 1:1 ratio expected from a supposed X-recessive gene ($\chi^2 = 7·12$; $P < 0·01$) and bears a resemblance to the 5:1 population ratio.

(iii) *Siblings*.—The available data on affected sib pairs also seem to point against an X-borne recessive gene.

McKeown and others (1951b) found among later sibs of propositi that 12/122 boys and 2/119 girls had the disease. Expected hypothetical $\dot{X}Y$ and $\dot{X}\dot{X}$ genotype frequencies are $(1+g)/2$ boys and g girls (Knox, 1958b). When $g = 0·2$, the expected sex ratio among affected subsequent sibs of a series of propositi is 3:1 instead of the population ratio of 5:1. The facts do not differ significantly from the expectation for sex-linkage, but resemble the population ratio more closely.

There are no other large-scale systematic studies of the disease in sibs, but a number of other sib-pairs are reported. From nine papers with such information (Brendle, 1937; Donovan, 1938; Adlington, 1943; Cockayne and Penrose, 1943; Shaefer and Erbes, 1948; Todd, 1949; McKeown and others, 1951b; Fenwick, 1953; Anon., 1958) the total of affected sib pairs was 57, distributed as:

MM, 40; MF*, 15; FF, 2.

It can be shown that for an X-recessive gene with equal penetrance in the two sexes the expected ratio MF*/FF is 2:1 at any gene frequency, and that the ratios MM/MF*/FF are $1:2g:g$ (Knox, 1958b). The expected distribution for 57 pairs, for a disease determined by an X-recessive gene frequency 0·2, is:

MM, 35·6; MF*, 14·2; FF, 7·1.

The expected distribution based upon a simple binomial distribution of the population sex ratio is:

MM, 36·5; MF*, 18·2; FF, 2·3.

Larger numbers would be needed to distinguish clearly between the two, and the difference between the observed 15:2 MF*/FF ratio and the expected 2:1 ratio in sex linkage is not significant ($\chi^2 = 3·65$; $0·05 < P < 0·1$). But it cannot be denied that the facts correspond more closely with expectations based upon less specific suppositions.

* Irrespective of order, *i.e.* MF+FM

(*iv*) *Relatives as a Whole*.—While the evidence so far presented provides a somewhat paradoxical pattern, one very clear point emerges when the frequencies of disease in different relatives are compared. This is that the full facts of familial concentration cannot be explained in terms of the supposed pattern of inheritance, and that a sex-linkage hypothesis would require considerable supplementation if it were to be comprehensive.

The evidence for this is as follows:

Subject to the possibility of inaccuracy from unequal penetrance in the two sexes, we saw that the frequency of the gene must be about 0·2 and the penetrance about 1/40. It was also shown that the supposed penetrance in cousins appeared to be of the same order. But when the genotype frequencies are considered quantitatively in offspring and siblings the facts require us to suppose much higher levels of penetrance.

McKeown and MacMahon (1955) estimated the incidence of the disease in offspring of propositi to be about 69/1,000. The appropriate hypothetical genotype frequency in offspring of male propositi is g both for boys and girls, and the same factor applies to the daughters of female propositi. All the sons of female propositi are susceptible. Combining these results in correct proportions and substituting $g = 0·2$, we expect to find the appropriate genotypes in a little more than a quarter (4/15) of all offspring. Comparison with the observed incidence gives a penetrance factor greater than 1/4, ten times that in the population at large or in cousins.

McKeown and others (1951b) also found an incidence among subsequent sibs of propositi equal to 58/1,000, and the expected frequency of the appropriate genotypes among these sibs is 0·4 (Knox, 1958b). The penetrance factor here is higher than 1/7, again very different from the population and cousin factors.

It is also obvious that inaccuracy of the gene-frequency estimate used in these calculations cannot explain the differences between the high penetrance in offspring and siblings and the much lower penetrance in cousins and in the general population. A lower estimate of g gives an expectation of a distinct difference between the incidence in cousins and in the general population. A higher estimate of g requires that we suppose even wider differences between the penetrance factors in sibs and offspring and in the population.

The similarity between the incidence in cousins and in the population has a more general implication. It limits the degree of genetic specificity which we can legitimately suppose in affected children.

Such limitation applies not only to the lower ranges of frequency of a hypothetical sex-linked recessive gene but also to specification in terms of substitutions at more than one locus. This conclusion applies to autosomal as well as sex-linked transmission, and means, in effect, that the main causes of familial concentration are almost certainly non-genetic. It probably applies also to maternal-foetal interactions of the type seen in haemolytic disease of the newborn, because a complete analogy with this disease demands genetic specification of the mother as well as the child. The falling incidence with birth rank, the opposite of the haemolytic disease pattern is also against a simple explanation in this class.

The appeal of the sex-linkage hypothesis therefore is on grounds other than the facts of familial concentration and these facts must be explained in some other way.

SELECTION AND GENE-FREQUENCY MAINTENANCE

The selection and evolutionary aspects of this disease offer some evidence on the plausibility of the sex-linkage hypothesis.

One difficulty in attempting to attribute any common serious disease to a harmful gene is the problem of how a stable gene frequency has been maintained in the face of strong selection. If we suppose that, until recent times genetically speaking, pyloric stenosis was always a fatal disease, there would be a loss of seven hypothetical X-borne harmful genes for every six cases (one ẌẌ and five ẌY). This is the incidence in 2,000 births or 3,000 X-chromosomes. Because we suppose that 4/5 of the 3,000 chromosomes are normal, this represents a loss of seven abnormal per 2,400 normal genes or 1/340 per generation, and if maintenance were by mutation would imply a mutation rate of the same order. This is many times greater than the accepted general level of mutation rates, namely about 1/100,000.

It is in relation to this discrepancy that the phenomenon of a disturbed sex ratio in the parental sibships has particular interest with respect to the possibility that the gene frequency is maintained by female (XẌ) heterosis. There are more female sibs than male sibs in each of three groups of parental sibships from which the propositi drew X-chromosomes, but not in the fourth group, the sibships of fathers of affected boys (see Table, opposite).

But, if this is indeed an expression of heterosis, there is reason for surprise at its extent. There were 1,048 aunts and 911 uncles in the three parental

TABLE

UNCLES AND AUNTS OF 445 BOYS AND GIRLS WITH PYLORIC STENOSIS

Sex of Propositus	Paternal		Maternal	
	Uncles	Aunts	Uncles	Aunts
Affected Boys (369)	648 (1·76)	641 (1·74)	683 (1·85)	739 (2·00)
Affected Girls (76)	114 (1·50)	150 (2·00)	114 (1·50)	159 (2·09)
Total 445	762	791	797	898

Mean number per child in parentheses

sibship groups with a disturbed ratio. The genotype frequencies expected among these aunts on the basis of a gene so frequent as 0·2 are:

$$XX, 0·19; X\dot{X}, 0·66; \dot{X}\dot{X}, 0·15$$

compared with population frequencies

$$XX, 0·64; X\dot{X}, 0·32; \dot{X}\dot{X}, 0·04$$

(Knox, 1958b).

By taking the survival value of the genotype XX as 1·0, the value for X$\dot{X}$ as 0·975 (1/40 lost), and by presuming that the excess of 1,048 over 911 is entirely due to X$\dot{X}$ heterosis, the survival factor for this last genotype comes to 1·56. If the X$\dot{X}$ survival value is less than the arbitrary 0·975, the survival factor for X$\dot{X}$ is higher still, as high as 1·91 when the $\dot{X}\dot{X}$ factor reaches zero. A survival factor of 1·56 for X$\dot{X}$ females applied to the total population results in a gain of about one $\dot{X}$ gene per thirteen normal per generation, many times the loss which could be attributed to the disease itself. It is possible to mitigate some of this imbalance by postulating other losses of $\dot{X}$ genes in the form of unrecognized loss of $\dot{X}\dot{X}$ females, but even with zero survival of $\dot{X}\dot{X}$ (and increase of the X$\dot{X}$ survival factor to 1·91) the expected net gain per generation is still two-thirds of that already given.

The supposition that the sex ratio disturbance in parents' sibs is entirely due to a lethal effect in $\dot{X}$Y males results in equally serious difficulties through aggravating the question of how the gene frequency is maintained.

This results in a situation where we must either accept both hypotheses, loss of $\dot{X}$Y uncles and heterosis in X$\dot{X}$ aunts, or accept neither. While we cannot regard this as a formal *reductio ad absurdum*, there are now a sufficient number of interdependent variables to make the acceptance of neither the natural preference.

This preference is readily supported by an examination of the sex ratios of cousins and later sibs. In these relatives, as in the parents' sibs, we expect frequencies of the X$\dot{X}$ and $\dot{X}$Y genotypes to be

higher than in the population (Knox, 1958b). Yet the published data show normal sex ratios (McKeown and others, 1951b). It seems improbable that secular changes sufficient to account for the disappearance of so striking a phenomenon could have occurred in the course of one generation and less likely that parents could have forgotten their male sibs dying relatively late in life on a scale which could explain the difference.

As with the phenomenon of familial concentration of the disease, so with the localization of the disturbed sex ratio in parents' sibs; the causes are to be sought largely or entirely among mechanisms other than sex-linked inheritance.

CONCLUSIONS

The conclusions arising from the above considerations are essentially negative.

The economy of the hypothesis of sex-linkage in explaining both the sex ratio of the disease and the distribution of the disturbed sex ratio in the parents' sibs was supported by the sex ratio observed among affected parents of affected sibs and asymmetry of the attack rates in maternal and paternal cousins. But other features of the disease, including a paucity of FF sib pairs and an unexpected sex ratio among affected children of affected fathers, pointed against this.

The hypothesis was impaired further by the demonstration that it failed to account for all the facts of the familial concentration of the disease and that its validity depended upon finding other means of explaining the greater part of the disparities between the attack rates in parents, sibs, and offspring of propositi, and attack rates in cousins. Such a demonstration cannot serve as formal evidence that there is no genetic determinant, for any genetic hypothesis invoking the concept of limited penetrance must allow the possibility of familial concentrations of environmental factors. But because it is necessary in any case to invoke non-genetic factors in explanation of the facts, it will be accepted that these same facts cannot serve as evidence that genetic factors do exist.

Consideration of selective effects and of gene frequency maintenance added further difficulties to the hypothesis, since it is improbable that the gene could be maintained at a stable level by mutation. Although it seemed at first that the sex ratio disturbance in the parents' sibs might supply a mechanism for such maintenance, critical examination showed that unless a very complex scheme were adopted the disturbance in fact aggravated the problem. Indeed, such difficulties were added that it became difficult to

accept that the disturbance in the parents' sibs could have any simple genetic basis.

The evidence as a whole, therefore, does not permit acceptance of the initial hypothesis. Although it is a negative conclusion the reasoning behind it has been presented in detail because of its importance in the context from which the hypothesis was drawn. A sex-linked genetic explanation of the phenomenon of disturbed ratios among the parents' sibs would have suggested simple genetic explanations for analogous phenomena in other congenital diseases (Knox, 1958b; Stewart, Ferguson-Smith, Lennox and Mack, 1958). Instead, it now seems necessary to replace pyloric stenosis in a class with those other diseases for whose curious ancestor-patterns we likewise have no ready explanation, and to consider other causes.

REFERENCES

Adlington, J. L. (1949). *Brit. med. J.*, 2, 740.
Anon. (1958). *Ibid.*, 1, 117.
Brendle, E. (1937). *Mschr. Kinderheilk.*, 68, 295.
Carter, C. O., and Powell, B. W. (1954). *Lancet*, 1, 746.
—— and Savage, T. R. (1951). *Arch. Dis. Childh.*, 26, 50.
Cockayne, E. A. (1934). *Lancet*, 1, 898.
—— (1938). *Arch. Dis. Childh.*, 13, 249.
—— and Penrose, L. S. (1943). *Ohio J. Sci.*, 43, 1.
Donovan, E. J. (1938). *Amer. J. Surg.*, 39, 377.
Fenwick, T. (1953). *Brit. med. J.*, 2, 12.
Knox, G. (1958a). (In press).
—— (1958b). *Brit. J. prev. soc. Med.*, 12, 183
McKeown, T., and MacMahon, B. (1955). *Arch. Dis. Childh.*, 30, 497.
——, ——, and Record, R. G. (1951a). *Ann. Eugen. (Lond.)*, 16, 249.
——, ——, —— (1951b). *Ibid.*, 16, 260.
——, ——, —— (1952). *Arch. Dis. Childh.*, 27, 386.
Nielsen, O. S. (1954). *Acta paediat. (Uppsala)*, 43, 522.
Shaefer, A. A., and Erbes, J. (1948). *Surg. Gynec. Obstet.*, 86, 45.
Stewart, J. S. S., Ferguson-Smith, M. A., Lennox, B., and Mack, W. S. (1958). *Lancet*, 2, 117.
Todd, R. McLaren (1949). *Brit. med. J.*, 2, 931.

APPENDIX

EXPECTED FREQUENCY OF GENOTYPES $\dot{X}X$ AND $\dot{X}Y$ AMONG COUSINS, SUPPOSING THAT PYLORIC STENOSIS IS CONDITIONALLY DETERMINED BY THESE GENOTYPES

Type of Cousin		Frequencies				
		General		For $g=0.2$		
		$\dot{X}Y$	$\dot{X}X$	$\dot{X}Y$	$\dot{X}X$	$\dot{X}X+\dot{X}Y$
Of Maternal	Uncles	$\dfrac{g}{(3+5g)/8}$	$g(1+3g)/4$	1/5	2/25	0·140
	Aunts		$g(3+5g)/8$	1/2	1/10	0·300 } 0·2200
Of Paternal	Uncles	$\dfrac{g}{g(5+3g)/4 (1+g)}$	$g^2(3+g)/2(1+g)$	1/5	4/75	0·127
	Aunts		$g^2(5+3g)/4(1+g)$	7/30	7/150	0·140 } 0·1335
General Population		g	g^2	1/5	1/25	0·120

The absolute genotype frequencies of 0·2200 and 0·1335 in maternal and paternal cousins are equivalent to proportional frequencies of 0·622 and 0·378. The mean of 0·2200 and 0·1335 is 0·177 compared with a population frequency of 0·120 and gives an expected attack rate in cousins of 4·4 per 1,000 compared with a population attack rate of 3 per 1,000. (See also Knox 1958b.)