Ann. Hum. Genet., Lond. (1963), **27**, 41
Printed in Great Britain

Some possible effects of birth control on the incidence of disorders and on the influence of birth order

By LEO A. GOODMAN*
University of Chicago

INTRODUCTION

This paper is concerned with the following two questions: (I) If the birth of a child with a particular disorder were to deter the parents from further reproduction, what (if any) effect would this have on the relative incidence of the disorder? (II) If family size were selectively limited in this way, what (if any) effect would this have on the relation between birth order and the relative incidence of the disorder?

Question (I) was discussed earlier by Lancelot Hogben (1952), and his results would suggest that birth control used to deter further reproduction when a child with a particular disorder is born would have no effect on the relative incidence of the disorder. [Hogben (1952) stated that the expected proportion of affected sibs obtained if selective limitation were practised would be equal to the expected proportion of affected sibs in a 'complete pool of...fraternities chosen randomwise' in the situation where there is no selective limitation.] In the present paper, we shall show that, under certain circumstances, selective limitation will actually reduce the relative incidence of affected individuals in the population. We find that this reduction can be non-trivial in magnitude. [It is, of course, obvious that selective limitation will tend to reduce the expected total number of offspring (both affected and unaffected) in the population, but the effect of this on the *relative* incidence of affected individuals is not so immediately obvious.]

With respect to question (II), we must consider separately the two kinds of procedures usually used to study the relation between birth order and the relative incidence of a disorder; viz. (*a*) procedures that depend on the formation of a 'control group' built up from the brothers and sisters of affected cases by suitable statistical techniques (e.g. the reconstruction suggested by Greenwood & Yule, 1914, or the modification of this technique suggested by Penrose, 1934), and (*b*) procedures that depend on a comparison of affected cases with a control group representative of all births. Penrose (1954) has noted that when procedures of type (*a*) are used 'there is a possibility of errors being caused by voluntary family limitation'. Smith & Record (1955), commenting on the two types of procedures (*a*) and (*b*), state that use of type (*a*) 'may give a false result if the birth of an affected child deters the parents from further reproduction', but that 'the difficulty does not arise' when type (*b*) is used. In the present paper, we shall show that the relation between birth order and the relative incidence of a disorder will be affected by the selective limitation of family size when procedures of type (*b*) are used as well as when procedures of type (*a*) are used. We shall find that, under certain circumstances, the influence of being first-born will be exaggerated in a positive direction when a control group representative of all births is used (i.e. when a type (*b*) procedure is used), and in a negative direction when the Greenwood–Yule method or the Penrose method are used (i.e. when type (*a*) procedures are used). Thus, if

* This research was supported in part by Research Grant No. NSF-G10368 from the Division of Mathematical, Physical and Engineering Sciences of the National Science Foundation.

the relative incidence of a disorder is higher among first-born than among later births when a control group representative of all births is used to compute the relative incidences (see, e.g. Lenz, 1959), this apparent 'positive effect of being first-born' may be due (in part) simply to selective limitation rather than to a real birth-order effect; and if the relative incidence of a disorder is the same for the different birth ranks, there may be, in fact, a real 'negative effect of being first-born' which has been obscured by the practice of selective limitation. Similarly, if the incidence of a disorder among first-born is lower than the corresponding expected value computed by the Greenwood–Yule method or the Penrose method, this apparent 'negative effect of being first-born' may be due (in part) simply to selective limitation rather than to a real birth-order effect; and if the incidence of a disorder among first-born is equal to the corresponding expected value (see, e.g. Boyer, Ferguson-Smith & Grumbach, 1961), there may actually be a real 'positive effect of being first-born' which has been obscured by selective limitation.

The results presented herein are related to an earlier article by the present author on the effects of birth control on the sex ratio (Goodman, 1961). In the earlier article we find, among other things, that the desire to produce a boy (a male heir), where the parents continue to bear children until a boy is born and then stop, would tend (under certain circumstances) to reduce the relative proportion of boys in the population. The present results would suggest that the reduction of the relative proportion of boys in the population would occur also under somewhat more general circumstances than studied earlier and that this kind of selective limitation of family size, where the birth of a boy would deter the parents from further reproduction, would also have an effect upon the relation between the sex ratio and birth order. The apparent relationship between the sex ratio and birth order would be exaggerated in a negative direction when there is this kind of selective limitation. Thus, the well-known negative 'birth-order effect' on the human sex ratio, discussed by many authors (see, e.g. Meyers, 1941; Lowe & McKeown, 1950; Novitski & Sandler, 1956), may be due (in part) to this selective limitation with respect to the sex of the offspring.

The formulae published earlier on the effects of selective limitation with respect to the sex of the offspring (Goodman, 1961) cannot be applied directly to the study of the effects of selective limitation with respect to a disorder, but the formulae presented herein, which are in a sense generalizations of the earlier formulae, can be applied to the study of the effects of selective limitation with respect to any characteristic, with respect to sex, with respect to disorders, etc. However, in the present paper, we shall for the sake of simplicity interpret the formulae given only for the case where there is selective limitation with respect to a disorder. A separate report by the author will discuss in detail the effects of selective limitation with respect to sex on the relation between the sex ratio and birth order and on other relationships.

In the present paper, we consider the case where the birth of one child with a particular disorder would deter the parents from further reproduction, and also the case where the birth of two affected offspring would deter further reproduction. In the former case we say that the criterion of termination is $A = 1$, and in the latter case say that it is $A = 2$. Focusing attention on the set of children (i.e. the sibship) born to parental pairs, we shall initially postulate (following Hogben, 1952) a fixed value of s, the size sibships would attain if there were no affected offspring. We shall consider all sets of parents in the population or in a sample from it. Numbering them from $i = 1$ to N, we shall initially postulate a fixed value of p_i, the probability that a child born to the ith set of parents will be affected ($i = 1, 2, \ldots, N$), and we assume that, with respect to any given birth, p_i does not depend upon preceding births. [We shall not assume as Hogben (1952) did (implicitly) that for the population of parents he was considering $p_i = p$ for all i; i.e. that all parents had the same probability of giving birth to an affected offspring.] The assumptions that we make here initially are for the sake of simplicity. Later in this article, we shall present more general results that remain valid even when some of these assumptions do not hold.

THE RELATIVE INCIDENCE OF A PARTICULAR DISORDER

Let $R = M/T$, where M is the expected number of affected children born, and T is the expected total number of children born. [Note that, strictly speaking, R is not the expected value of the proportion of affected sibs observed among the N families, but it will be an approximation to this expected proportion when the number of sibships considered is a large sample from the population of interest (assumed infinite).] We shall now compute the value of R, and shall show that R is always less when there is selective limitation (when the criterion of termination is $A = 1$ or $A = 2$) than when there is no selective limitation of family size (provided that the values of p_i $(i = 1, 2, \ldots, N)$ are not all equal to each other for the population under consideration).

Let P_i $(X \geqslant t)$ denote the probability that the ith set of parents will have t or more children. When $A = 1$, then

$$T = \sum_{i=1}^{N} \sum_{t=1}^{s} P_i(X \geqslant t) = \sum_{i=1}^{N} \sum_{t=1}^{s} q_i^{t-1} = \sum_{i=1}^{N} (1-q_i^s)/p_i,$$

$$M = \sum_{i=1}^{N} \sum_{t=1}^{s} p_i q_i^{t-1} = \sum_{i=1}^{N} (1-q_i^s),$$

$$R = \sum_{i=1}^{N} (1-q_i^s) \Big/ \left[\sum_{i=1}^{N} (1-q_i^s)/p_i\right]. \tag{1}$$

Writing $\sum_{i=1}^{N} f(i)/N = E\{f\}$, we note that

$$R = E\{1-q^s\}/E\{(1-q^s)/p\}. \tag{2}$$

Since $(1-q^s)/p$ is a decreasing function of p (when $s \geqslant 2$, $0 < p \leqslant 1$), the covariance between p and $(1-q^s)/p$ will be negative, which implies that

$$E\{[(1-q^s)/p][p]\} < E\{(1-q^s)/p\}E\{p\},$$

and that $R < E\{p\}$. This inequality is replaced by equality only if $p_i = p$ for all i. Since $E\{p\}$ is the value of R when sibship size is not selectively limited (assuming s is a fixed constant), we find that the value of R is reduced when $A = 1$.

When $p_i = p$ for $i = 1, 2, \ldots, N$, we find from (1) or (2) that $R = p$, which is in agreement with the result given by Hogben (1952) for this special case. When an upper bound has not been set for sibship size (i.e. when $s = \infty$), we find from (2) that $R = 1/E\{1/p\}$, which is in agreement with the result given by Goodman (1961).

When $A = 2$, then

$$T = \sum_{i=1}^{N} \sum_{t=1}^{s} P_i(X \geqslant t) = \sum_{i=1}^{N} \left\{1 + \sum_{t=2}^{s} [q_i^{t-1} + (t-1) p_i q_i^{t-2}]\right\}$$

$$= \sum_{i=1}^{N} \{[2(1-q_i^s)/p_i] - s q_i^{s-1}\},$$

$$M = \sum_{i=1}^{N} \sum_{t=1}^{s} p_i P_i(X \geqslant t) = \sum_{i=1}^{N} \{2(1-q_i^s) - s p_i q_i^{s-1}\},$$

$$R = \sum_{i=1}^{N} \{2(1-q_i^s) - s p_i q_i^{s-1}\} \Big/ \sum_{i=1}^{N} \{[2(1-q_i^s)/p_i] - s q_i^{s-1}\}. \tag{3}$$

Thus $$R = E\{2-sq^{s-1}+(s-2)\,q^s\}/E\{[2-sq^{s-1}+(s-2)\,q^s]/p\}. \tag{4}$$

Since $[2-sq^{s-1}+(s-2)\,q^s]/p$ is a decreasing function of p (when $s \geqslant 3$, $0 < p \leqslant 1$), the covariance between p and $[2-sq^{s-1}+(s-2)\,q^s]/p$ will be negative, which implies that

$$E\{[2-sq^{s-1}+(s-2)\,q^s]\} < E\{[2-sq^{s-1}+(s-2)\,q^s]/p\}\,E\{p\},$$

and that $R < E\{p\}$. (Here too the inequality is replaced by equality only if $p_i = p$ for all i.) Thus, the value of R is smaller when $A = 2$ than when there is no selective limitation.

Having proved that $R < E\{p\}$ for $A = 1$ and $A = 2$, it is also possible to prove the more general result that the value of R will be less with any criterion of termination of the kind considered here (with $A < s$) than with no selective limitation at all, but we shall not go into these details here.

For purposes of illustrating the application of (2) and (4), we shall first examine the extreme (perhaps unrealistic) situation where the disorder under consideration is such that $p_i = 1$ for $\epsilon(100)\,\%$ of the population, and $p_i = 0$ for $(1-\epsilon)\,(100)\,\%$ of the population $(0 < \epsilon \leqslant 1)$. Then when $A = 1$, we calculate from (2) that

$$R = \epsilon/[\epsilon+s(1-\epsilon)],$$

and $$R/E\{p\} = 1/[s-\epsilon(s-1)]. \tag{5}$$

When $A = 2$, we calculate from (4) that

$$R = 2\epsilon/[2\epsilon+s(1-\epsilon)]$$

and $$R/E\{p\} = 2/[s-\epsilon(s-2)]. \tag{6}$$

Thus, when the disorder is rather infrequent (i.e. when the value of ϵ is near zero), the relative decrease in R would be approximately $(s-1)/s$ with $A = 1$ and $(s-2)/s$ with $A = 2$. For example, if $\epsilon = 0{\cdot}01$ then the disorder would occur in $1/(100)$ of the births when there is no selective limitation (i.e. $E\{p\} = 0{\cdot}01$), the value of R would be $1/(298)$ for $A = 1$ and $1/(149{\cdot}5)$ for $A = 2$ when $s = 3$, and $1/(397)$ for $A = 1$ and $1/(199)$ for $A = 2$ when $s = 4$; if $\epsilon = 0{\cdot}001$, then $R = 1/(1000)$ when there is no selective limitation, and the value of R would be $1/(2998)$ for $A = 1$ and $1/(1499{\cdot}5)$ for $A = 2$ when $s = 3$, and $1/(3997)$ for $A = 1$ and $1/(1999)$ for $A = 2$ when $s = 4$. In these cases, the relative decrease in R is approximately $2/3$ for $A = 1$ and $1/3$ for $A = 2$ when $s = 3$, and $3/4$ for $A = 1$ and $1/2$ for $A = 2$ when $s = 4$. Though the relative decrease in R will be somewhat less pronounced when the disorder is not so infrequent, the reduction will nevertheless remain non-trivial (when, say, $0 < \epsilon \leqslant 1/2$).

In passing, we note that in the situation considered here the relative decrease in the expected number of affected sibs will be $(s-1)/s$ for $A = 1$ and $(s-2)/s$ for $A = 2$ regardless of whether the disorder is more or less infrequent. In other words, for all values of ϵ $(0 < \epsilon \leqslant 1)$ the relative decrease in the expected number of affected sibs will be as stated above.

From (5) and (6) we note that the relative decrease in R is greater with $A = 1$ than with $A = 2$ for any fixed value of s, that with $A = 1$ and with $A = 2$ the reduction in R increases with increasing values of s, and that the reduction in R attained with $A = 1$ for a fixed value of $s = s_0$ can be attained with $A = 2$ and $s = 2s_0$ in the special case considered here.

Let us next examine the situation where the disorder under consideration is such that $p_i = p_1$ for $\epsilon(100)\,\%$ of the population and $p_i = 0$ for $(1-\epsilon)\,(100)\,\%$ of the population. The value of p_1 might be, say $1/4$, or more generally $0 < p_1 \leqslant 1$. For $A = 1$,

$$R/E\{p\} = 1/[\epsilon+sp_1(1-\epsilon)/(1-q_1^s)], \tag{7}$$

and for $A = 2$ $$R/E\{p\} = 1/\{\epsilon+sp_1(1-\epsilon)/[2-sq_1^{s-1}+(s-2)\,q_1^s]\}. \tag{8}$$

Formulas (7) and (8) are generalizations of (5) and (6), respectively. From (7) and (8), we note that when the disorder is rather infrequent (i.e. when ϵ is near zero), then $R/E\{p\}$ is approximately $(1-q_1^s)/sp_1$ for $A = 1$ and $[2-sq_1^{s-1}+(s-2)\,q_1^s]/sp_1$ for $A = 2$. For both $A = 1$ and $A = 2$, the value of $R/E\{p\}$ decreases with increasing values of p_1.

We now consider the situation where the disorder is such that $p_i = p_1$ for $\epsilon(100)\,\%$ of the population and $p_i = p_0$ for $(1-\epsilon)\,(100)\,\%$ of the population, where $0 < p_0 \leqslant p_1 \leqslant 1$. By applying (2) and (4), the value of $R/E\{p\}$ can be calculated directly for $A = 1$ and for $A = 2$. The following approximations can be applied when ϵ is near zero: For $A = 1$

$$R/E\{p\} \approx 1-\{\epsilon[(p_1/p_0)-1]\,[1-[p_0(1-q_1^s)/p_1(1-q_0^s)]]\}, \tag{9}$$

and for $A = 2$

$$R/E\{p\} \approx 1-\left\{\epsilon[(p_1/p_0)-1]\left[1-\left[\frac{(2-sq_1^{s-1}+(s-2)\,q_1^s)\,p_0}{(2-sq_0^{s-1}+(s-2)\,q_0^s)\,p_1}\right]\right]\right\}. \tag{10}$$

When s is large, these approximations can be simplified to

$$R/E\{p\} \approx 1-\epsilon(p_1-p_0)^2/p_0p_1,$$

which indicates that the relative decrease in R will be approximately $\epsilon(p_1-p_0)^2/p_0p_1$ in the situation considered here.

Though the situations considered above are perhaps somewhat unrealistic, they do provide simple illustrations of (*a*) the application of the more general formulas (2) and (4), and (*b*) certain more general qualitative relations concerning the possible range of effect on R of different values of A. For any distribution of values of p_i (e.g. where the p_i may take on any number of different possible values, such as $\frac{1}{4}$, $\frac{1}{2}$, etc.), the formulas (2) and (4) can be applied directly to calculate the reduction in R when $A = 1$ and when $A = 2$. By applying (2) and (4) it is also possible to prove that for any distribution of p_i (where the p_i are not all equal to each other) the value of R will be less with $A = 1$ than with $A = 2$ for any fixed value of s, and that it will decrease for both $A = 1$ and $A = 2$ with increasing values of s. More generally, for any distribution of values of p_i (where the p_i are not all identical), and for any criterion of termination of the kind discussed here (i.e. for any value of $A < s$), the value of R will decrease with increasing values of s, and for any fixed value of s the value of R will decrease with decreasing values of A.

For any given set of parents (say, the ith set) let $R_i = M_i/T_i$, where M_i is the expected value of the number m_i of affected sibs in the ith family and T_i is the expected value of the total number t_i of sibs in that family. For this family, the value of R_i will remain unchanged no matter what kind of termination criteria it adopts, assuming that the criteria are such that $T_i < \infty$. (See related comments by Goodman, 1961.) The value of R_i will remain equal to p_i. Since T_i is a decreasing function of p_i for both $A = 1$ and $A = 2$ (when $A < s$), and since

$$R = \sum_{i=1}^{N} R_i T_i \Big/ \sum_{i=1}^{N} T_i = \sum_{i=1}^{N} p_i T_i \Big/ \sum_{i=1}^{N} T_i$$

(i.e. R is a weighted average of the p_i that gives the smaller values of p_i relatively more weight) we found earlier herein that $R \leqslant E\{p\}$ when $A = 1$ or $A = 2$, and that $R = E\{p\}$ only if all p_i are identical (for $i = 1, 2, \ldots, N$).

Though $R_i = p_i$ for both $A = 1$ and $A = 2$, the expected value ρ_i of m_i/t_i will usually differ from p_i. (Where there is no selective limitation and the value of t_i is fixed, then $\rho_i = p_i$.) For any value of A, the criterion of termination, it is well known (see, e.g. Kendall & Stuart, 1958) that $(m_i-1)/(t_i-1)$ is an unbiased estimate of p_i when $s = \infty$, which indicates that (when $p_i < 1$) in this particular case $\rho_i > p_i$, since $(m_i - 1)/(t_i - 1) \leqslant m_i/t_i$. Thus, the criterion of termination applied by the ith set of parents will affect the value of ρ_i (though not the value of R_i), and as we observed earlier it will also have its effect on R. (We note in passing that, for any value of A, the expected value γ_i of t_i/m_i will be equal to $1/p_i$ when $s = \infty$, but that γ_i will usually differ from $1/p_i$, except when the value of m_i is fixed.)

THE RELATION BETWEEN RANK ORDER AND THE RELATIVE INCIDENCE OF A PARTICULAR DISORDER

Let $R^{(j)} = M^{(j)}/T^{(j)}$, where $M^{(j)}$ is the expected number of affected children of the jth birth rank, and $T^{(j)}$ is the expected total number of children of the jth birth rank. We note that

$$R = \sum_{j=1}^{s} M^{(j)} \Bigg/ \sum_{j=1}^{s} T^{(j)}.$$

We have $R^{(j)} = E\{p\}$ for $j = 1, 2, \ldots, s$ when there is no selective limitation of sibship size (assuming s fixed). With $A = 1$, we have $R^{(1)} = E\{p\}$ and

$$R^{(j)} = \sum_{i=1}^{N} p_i q_i^{j-1} \Bigg/ \sum_{i=1}^{N} q_i^{j-1} \tag{11}$$

for $j = 1, 2, \ldots, s$. With $A = 2$, we have $R^{(1)} = R^{(2)} = E\{p\}$ and

$$R^{(j)} = \sum_{i=1}^{N} p_i[q_i^{j-1} + (j-1)\, p_i q_i^{j-2}] \sum_{i=1}^{N} [q_i^{j-1} + (j-1)\, p_i q_i^{j-2}] \tag{12}$$

for $j = 1, 2, \ldots, s$. Thus, with $A = 1$

$$R^{(j)} = E\{pq^{j-1}\}/E\{q^{j-1}\}, \tag{13}$$

and with $A = 2$

$$R^{(j)} = E\{p[q^{j-1} + (j-1)\, pq^{j-2}]\}/E\{q^{j-1} + (j-1)\, pq^{j-2}\}. \tag{14}$$

For purposes of illustrating the application of (13) and (14), we first examine the extreme (unrealistic) situation where the disorder under consideration is such that $p_i = 1$ for $\epsilon(100)\,\%$ of the population and $p_i = 0$ for $(1-\epsilon)(100)\,\%$ of the population $(0 < \epsilon \leqslant 1)$. When $A = 1$, we calculate from (13) that $R^{(1)} = \epsilon$ and $R^{(j)} = 0$ for $j = 2, 3, \ldots, s$. When $A = 2$, we calculate from (14) that $R^{(1)} = R^{(2)} = \epsilon$, and $R^{(j)} = 0$ for $j = 3, 4, \ldots, s$.

When the disorder is such that $p_i = p_1$ for $\epsilon(100)\,\%$ of the population and $p_i = 0$ for $(1-\epsilon)(100)\,\%$ of the population $(0 < p_1 < 1)$, then from (13) we find that

$$R^{(j)}/E\{p\} = 1/[\epsilon + (1-\epsilon)/q_1^{j-1}] \tag{15}$$

for $A = 1$, and from (14)

$$R^{(j)}/E\{p\} = 1/\{\epsilon + (1-\epsilon)/[q_1^{j-1} + (j-1)\, p_1 q_1^{j-2}]\} \tag{16}$$

for $A = 2$. From (15) and (16) we note that when the disorder is rather infrequent (i.e. when ϵ is near zero), then $R^{(j)}/E\{p\}$ is approximately q_1^{j-1} for $A = 1$ and $[q_1^{j-1} + (j-1)\, p_1 q_1^{j-2}]$ for $A = 2$. The values of $R^{(j)}$ form a decreasing series for $j = 1, 2, \ldots, s$ with $A = 1$, and a non-increasing series with $A = 2$ (actually decreasing for $j = 2, 3, \ldots, s$). For each j $(j = 2, 3, \ldots, s)$, the value of $R^{(j)}$ is less with $A = 1$ than with $A = 2$. Thus we see that in a situation where there would be no relation between birth rank and the relative incidence of a disorder if there were no selective limitation, an apparent 'positive effect of being first-born' is observed (i.e. $R^{(1)} > R^{(j)}$ for $j > 1$), this influence appearing greater when $A = 1$ than when $A = 2$.

When the disorder is such that $p_i = p_1$ for $\epsilon(100)\,\%$ the population and $p_i = p_0$ for $(1-\epsilon)(100)\,\%$ of the population $(0 < p_0 \leqslant p_1 \leqslant 1)$, the value of $R^{(j)}/E\{p\}$ can be calculated directly from (13) and (14) for $A = 1$ and $A = 2$, respectively. The following approximations can be applied when ϵ is near zero. For $A = 1$

$$R^{(j)}/E\{p\} \approx 1 - \epsilon[(p_1/p_0) - 1]\,[1 - (q_1/q_0)^{j-1}], \tag{17}$$

and for $A = 2$

$$R^{(j)}/E\{p\} \approx 1 - \left\{\epsilon[(p_1/p_0) - 1]\left[1 - \left[\frac{q_1^{j-1} + (j-1)\,p_1 q_1^{j-2}}{q_0^{j-1} + (j-1)\,p_0 q_0^{j-2}}\right]\right]\right\}. \tag{18}$$

From these approximations, we obtain

$$R^{(2)}/E\{p\} \approx 1 - \epsilon(p_1 - p_0)^2/p_0 q_0$$

for $A = 1$, and

$$R^{(3)}/E\{p\} \approx 1 - \epsilon(p_1 - p_0)^2\,(p_1 + p_0)/p_0(1 - p_0^2)$$

for $A = 2$. When j is large, (17) and (18) can be simplified (for $j \leqslant s$) to

$$R^{(j)}/E\{p\} \approx 1 - \epsilon(p_1 - p_0)/p_0,$$

which indicates that the relative decrease in $R^{(j)}$ will be approximately $\epsilon(p_1 - p_0)/p_0$ in the situation considered here ($j \to \infty, s \to \infty$).

The situations considered above provide simple illustrations of (*a*) the application of (13) and (14), and (*b*) certain more general relations concerning the possible range of effect on $R^{(j)}$ of different values of A. For any distribution of p_i, the formulas (13) and (14) can be applied directly to calculate the reduction in $R^{(j)}$ when $A = 1$ and when $A = 2$. By applying (13) and (14), it is also possible to prove that for any distribution of p_i (where the p_i are not all equal, and $p_i q_i \neq 0$ for some i) the values of $R^{(j)}$ will be less with $A = 1$ than with $A = 2$ ($j = 2, 3, \ldots, s$), and that the $R^{(j)}$ will decrease for both $A = 1$ and $A = 2$ with increasing values of j (where $A \leqslant j \leqslant s$). More generally, for any criterion of termination A, the value of $R^{(j)}$ will decrease with increasing values of j (where $A \leqslant j \leqslant s$), and for any fixed value of j the value of $R^{(j)}$ will decrease with decreasing values of A. Thus we find (as we did earlier in the special case considered there) that, even if a control group representative of all births is used to study the relation between birth rank and the relative incidence of the disorder, in the situation described here where no relation would exist if there were no selective limitation, an apparent 'positive effect of being first-born' is observed when $A = 1$ or $A = 2$, this influence appearing greater with $A = 1$ than with $A = 2$.

Let us now consider the Greenwood–Yule method. Denoting the observed number of affected individuals from sibships of size x by f_x, they suggest that, in order to study the relation (if any) between birth rank and the relative incidence of the disorder, the observed number $m^{(j)}$ of affected individuals of the jth birth rank should be compared with $\sum_{x=j}^{\infty} (f_x/x)$, the 'number to be expected' based on the f_x assuming that there is no birth rank effect. When sibship size is not selectively limited, we might assume that, given that the ith sibship is of size x, the conditional probability is p_i that the kth individual in the sibship is affected ($k = 1, 2, \ldots, x$). In this case, the expected value ϕ_x of f_x is $x \sum_{i=1}^{N} p_i\,\theta_i(x)$, where $\theta_i(x)$ is the probability that the ith sibship is of size x, the expected value $\mu^{(j)}$ of $m^{(j)}$ is $\sum_{i=1}^{N} p_i\,\psi_i(j)$, where $\psi_i(j)$ is the probability that the ith sibship is of size j or more, and $\sum_{x=j}^{\infty} (\phi_x/x) = \mu^{(j)}$. (Where sibship size, say s, is a fixed value, then $\theta_i(x)$ will be 1 for $x = s$ and 0 otherwise.) Thus, the expected value of $m^{(j)}$ will equal the expected value of $\sum_{x=j}^{\infty} (f_x/x)$, which indicates that use of the Greenwood–Yule method when there is no selective limitation would usually lead in the situation considered here (ignoring sampling fluctuations) to the correct conclusion; viz. that the birth rank effect is nil in the situation under consideration.

When $A = 1$, then $f_x = m^{(x)}$ and $\phi_x = \sum_{i=1}^{N} \theta_i(x) = \mu^{(x)}$. Thus, when $f_x > 0$ for some $x > 1$, $m^{(1)}$ will be less than 'the number to be expected' $\sum_{x=1}^{\infty} (f_x/x)$, which indicates that use of the Greenwood–Yule method when $A = 1$ would usually lead in the situation considered here to the conclusion that there is a 'negative effect of being first-born' ($R^{(1)} < R$), though the birth rank effect would actually be nil if there were no selective limitation. These remarks relating to the Greenwood–Yule method can be applied as well to the modification of it proposed by Penrose (1934).

DISCUSSION

Following Hogben (1952), we postulated herein a fixed value of s, the size sibships would attain if there were no affected offspring. This is the size of sibships postulated for the situation where sibship size is not selectively limited. (In part of the preceding section we considered briefly the case where sibship size was a random variable x and showed there that the Greenwood–Yule method would yield correct results in this case (with selective limitation absent) as well as when s was fixed, provided that the conditional probability distribution of disorders for the ith sibship, given that this sibship's size was $x = s$, was the same as the probability distribution of disorders for this sibship when sibship size was fixed at s.) Hogben (1952) noted that his 'argument holds good for any value s may have; and the arbitrariness of the initial assumption that s has a fixed value does not invalidate it'. This will be so when $p_i = p$ for all i; but when the p_i need not all be equal to each other, conclusions that hold good for any value s may have will not necessarily hold good in general, say, when different sets of parents might have different s values. Denoting the value of s for the ith set of parents by s_i, we can divide the N sets of parents into mutually exclusive classes where all sets of parents in the xth class have $s_i = x$. For each class separately, the formulas and general relationships presented earlier herein can be applied directly since they hold good for each value of s. For all N sets of parents (where the value of s_i need not be identical for all i), many of the formulas presented earlier (where $s_i = s$ for all i) will require only slight modification, and many of the general relationships will continue to hold true when the p_i and s_i are independent. (When the p_i and s_i are not independent, the relationships presented earlier need no longer hold true in general.) We now describe the kinds of modifications that are required.

For $A = 1$, the value of R given by formula (1) when $s_i = s$ for all i can be generalized to cover the situation now under consideration by simply replacing s by s_i in (1). Formula (2) holds good for both the special case and the more general situation (where the value of s may be different for different sets of parents). For $A = 2$, formula (3) should be modified as (1) was, and (4) holds good as (2) did. For $A = 1$, the value of $R^{(j)}$ given by (11) can be generalized simply by replacing the summation over all N values of i by summation over those values of i such that $s_i \geqslant j$. For $A = 2$, a similar modification should be made in (12). Denoting by Σ^j the summation over those i such that $s_i \geqslant j$, by $N^{(j)}$ the number of values of i such that $s_i \geqslant j$, and by $E^{(j)}\{f\}$ the value of $\Sigma^j f(i)/N^{(j)}$, formulas (13) amd (14) can be generalized by simply replacing the symbols E by $E^{(j)}$

in these expressions. Applying these generalized formulas to the particular special cases discussed earlier herein will lead to straightforward modifications of the special formulas given there; e.g. of (7) and (8), (15) and (16), etc.

When p_i and s_i are independent, all formulas and relations concerning the $R^{(j)}$ presented in the preceding section (for $s_i = s$) will continue to hold good where now $R^{(j)}$ is defined for all j from 1 to the maximum s_i. Since $R = \sum_j T^{(j)}R^{(j)}/\sum_j T^{(j)}$, we note when the p_i and s_i are independent that $R \leqslant E\{p\}$, the strict inequality holding when the criterion of termination A is less than some s_i values and the p_i values are not all equal to each other. (If the p_i and s_i are not independent, R need not be lower when $A = 1$ than when there is no selective limitation; by applying the formulas given here, we can determine in any particular case whether or not this is so. The relationship can be reversed in some cases where high p_i is associated with low s_i.)

We have discussed here the effect of removing the restriction that $s_i = s$ for all i, and now we shall consider the effect of removing the restriction that the p_i are constant. We showed earlier that, if a control group representative of all births is used to study the birth rank effect in the situation where the p_i are constant (and $s_i = s$ for all i), when there is no selective limitation the effect is nil, but when there is selective limitation (when $A = 1$ or $A = 2$) the 'effect of being first-born' would appear to be positive. We shall now show that this exaggeration when $A = 1$ or $A = 2$ of the positive 'effect of being first-born' will continue to persist, under certain circumstances, even when the p_i are not constant.

Let p_{ij} denote the probability that the jth child born to the ith set of parents will be affected. Let $q_{ij} = 1 - p_{ij}$ and $Q^0_{ij} = \prod_{k=1}^{j-1} q_{ik}$. When sibship size is not selectively limited, we have (assuming s fixed)

$$R^{(j)} = \sum_{i=1}^{N} p_{ij}/N = E\{p_{.j}\}, \quad \text{for} \quad j = 1, 2, \ldots, s,$$

and when $A = 1$ we have

$$R^{(j)} = \sum_{i=1}^{N} p_{ij}Q^0_{ij} \Big/ \sum_{i=1}^{N} Q^0_{ij} = E\{p_{.j}Q^0_{.j}\}/E\{Q^0_{.j}\}, \tag{19}$$

where dots within the brackets indicate that a subscript i appeared there, the entire term within brackets having been averaged over all values of i from 1 to N. If $E\{p_{.1}\} > E\{p_{.j}\}$ (for example, if $p_{i1} \geqslant p_{ij}$, with strict inequality holding for some i) for $j = 2, 3, \ldots, s$, then there is a positive effect of being first-born when there is no selective limitation. With $A = 1$, the identical positive effect will appear if for each value of j ($j \geqslant 2$) there is no correlation between the p_{ij} and Q^0_{ij}, but if there is a negative correlation between these quantities (as there was between p_i and q_i^{j-1}) then

$$R^{(1)} = E\{p_{.1}\} \quad \text{and} \quad R^{(j)} < E\{p_{.j}\}$$

for $j \geqslant 2$, so that a comparison of $R^{(1)}$ with $R^{(j)}$ would exaggerate the positive effect of being first-born in this case. We note also that with $A = 1$ a negative effect of being first-born would be understated when the p_{ij} and Q^0_{ij} are negatively correlated for each j ($j \geqslant 2$).

Let Q^1_{ij} be the probability that there will be no more than one affected child among the first $(j-1)$ children of the ith set of parents; i.e.

$$Q^1_{ij} = \prod_{k=1}^{j-1} q_{ik}\left[1 + \sum_{r=1}^{j-1} (p_{ir}/q_{ir})\right].$$

Then with $A = 2$

$$R^{(j)} = \sum_{i=1}^{N} p_{ij}Q^1_{ij} \Big/ \sum_{i=1}^{N} Q^1_{ij}$$
$$= E\{p_{.j}Q^1_{.j}\}/E\{Q^1_{.j}\}. \tag{20}$$

If for each j ($j > 2$) the Q^1_{ij} are negatively correlated with the p_{ij} (as they were when $p_{ij} = p_i$ for $j = 1, 2, \ldots, s$), then $R^{(j)} < E\{p_{.j}\}$ for $j = 3, 4, \ldots, s$, which indicates that here too with $A = 2$ a positive 'effect of being first-born' will be exaggerated and a negative 'effect' will be understated.

We now consider briefly the Greenwood–Yule method in the situation where the p_i are not constant. As in the preceding section, we let $\mu^{(1)}$ denote the expected value of the number $m^{(1)}$ of affected first-born individuals, ϕ_x the expected value of the number f_x of affected individuals from sibships of size x, $\theta_i(x)$ the probability that the ith sibship is of size x. When sibship size is not selectively limited (and when the conditional probability distribution of disorders for the ith sibship, given that this sibship's size is $x = s$, is the same as the probability distribution of disorders for this sibship when the size is fixed at s), then $\mu^{(1)} = \sum_{i=1}^{N} p_{i1} = NE\{p_{.1}\}$ and the expected value of $\hat{m}^{(1)} = \sum_{x=1}^{\infty} (f_x/x)$ is

$$\begin{aligned}\hat{\mu}(1) &= \sum_{x=1}^{\infty} (\phi_x/x) = \sum_{i=1}^{N} \sum_{x=1}^{\infty} \sum_{j=1}^{x} \{[p_{ij}\,\theta_i(x)]/x\} \\ &= \sum_{i=1}^{N} \sum_{x=1}^{\infty} \bar{p}_{ix}\,\theta_i(x), \\ &= N \sum_{x=1}^{\infty} E\{\bar{p}_{.x}\,\theta_{.}(x)\},\end{aligned} \tag{21}$$

where $\bar{p}_{ix} = \sum_{j=1}^{x} p_{ij}/x$. When for each x there is no correlation between the $\bar{p}_{ix}$ and the $\theta_i(x)$, then

$$\hat{\mu}^{(1)} = N \sum_{x=1}^{\infty} E\{\bar{p}_{.x}\} E\{\theta_{.}(x)\}. \tag{22}$$

If $E\{\bar{p}_{.1}) > E\{\bar{p}_{.x}\}$ for $x \geqslant 2$, then from (22) we note that $\mu^{(1)} > \hat{\mu}^{(1)}$, which indicates that in the situation considered here (ignoring sampling fluctuations) the application of the Greenwood–Yule method would correctly detect the positive effect of being first-born. It should be noted, however, that a more complete assessment of the actual magnitude of this effect is (in principle) more readily available when the control group method is used (if there is no selective limitation) than when the Greenwood–Yule method is used. We also note that with $A = 1$ the 'effect of being first-born' will usually appear to be negative when the Greenwood–Yule method is applied, even in cases where the effect would be positive if there were no selective limitation.

Formulas presented earlier herein were concerned with the situation where all N sets of parents had the same criterion of termination; e.g. $A = 1$, $A = 2$ or $A = \infty$. We now consider the more general situation where the criterion of termination for the ith set of parents is A_i, and where the value of s for this set of parents is s_i. Let $Q^{B_i}_{ij}$ be the probability that the ith set of parents will produce no more than B_i affected children among their first $(j-1)$ children, where $Q^{B_i}_{ij} = 1$ for $j = 1$ and $Q^{B_i}_{ij} = 0$ for $j > s_i$. Then

$$R^{(j)} = \sum_{i=1}^{N} p_{ij} Q^{B_i}_{ij} \Big/ \sum_{i=1}^{N} Q^{B_i}_{ij} \tag{23}$$

and

$$R = \sum_{i=1}^{N} \sum_{j=1}^{s_i} p_{ij} Q^{B_i}_{ij} \Big/ \sum_{i=1}^{N} \sum_{j=1}^{s_i} Q^{B_i}_{ij}, \tag{24}$$

where $B_i = A_i - 1$. Formula (23) is a generalization of (11), (12), and formula (24) is a generalization of (1), (3). From (24) we see that if the correlation between the p_{ij} and $Q^{B_i}_{ij}$ is negative (where i ranges from 1 to N, and the j corresponding to a particular i value ranges from $j = 1$ to s_i), the value of R will be less when there is some selective limitation (i.e. when $A_i < s_i$ for some i) than when there is none.

We have observed here that many of the formulas and relations presented in the preceding sections can be modified to cover more general situations than those presented there. While many of the relations will remain true under these more general conditions, this will not always be the case. For example, applying (24) it is possible to show that the value of R can be less (or more) when $A = 1$ than when there is no selective limitation even if $p_{ij} = p_{kj}$ for all i, j, k, though we observed earlier that the value of R will be unaffected by selective limitation if $p_{ij} = p_{kj} = p$ for all i, j, k. Thus, the value of R will be affected by selective limitation not only when the disorder under consideration is such that $p_{ij} \neq p_{kj}$ for some i and k (e.g. when the disorder is genetic in character), but even when $p_{ij} = p_{kj}$ for all i and k (i.e. even when the chance that the jth offspring is affected is the same for all sets of parents), except when $p_{ij} = p_{kj} = p$ for all i, j, k.

SUMMARY

Selective limitation of family size will reduce the relative incidence R of a disorder if for different values of i there are differences in the probability p_i that a child born to the ith set of parents is affected, and if the size s_i of the ith sibship (in the absence of affected offspring) is independent of p_i. When $s_i = s$ for all i, R will be reduced more when the criterion of termination is $A = 1$ (i.e. when the presence of one affected offspring terminates the sibship) than when the value of A is greater than one, and this reduction in R for any value of A $(A < s)$ will increase with increasing values of s. To facilitate the measurement of R under various general conditions, and to assess the magnitude of the effect of selective limitation on R, various formulae are given herein.

The relation between birth rank and the relative incidence of a disorder will be affected by the selective limitation of family size regardless of whether the relation is determined by the control group method, by the Greenwood–Yule method, or by the Penrose (1934) modification of the Greenwood–Yule method. With the control group method (where the affected cases are compared with a control group representative of all births) the 'effect of being first-born' will be exaggerated in a positive direction when $A = 1$ if there is a negative correlation (for each j) between the probability p_{ij} that the jth offspring of the ith set of parents is affected and the probability Q^0_{ij} that there are no affected offspring among the first $(j-1)$ offspring of the ith set of parents. With the Greenwood–Yule method or the Penrose modification, the 'effect of being first-born' will be exaggerated in a negative direction when $A = 1$; even if this effect is positive when there is no selective limitation, it will appear to be negative when the Greenwood–Yule method is used with $A = 1$. To facilitate the measurement of the effect of selective limitation on the 'birth rank effect', various formulae are given herein.

REFERENCES

BOYER, S. H., FERGUSON-SMITH, M. A. & GRUMBACH, M. M. (1961). The lack of influence of parental age and birth order in the aetiology of nuclear sex chromatin-negative Turner's syndrome. *Ann. Hum. Genet., Lond.*, **25**, 215.

GOODMAN, L. A. (1961). Some possible effects of birth control on the human sex ratio. *Ann. Hum. Genet., Lond.*, **25**, 75.

GREENWOOD, M. & YULE, G. U. (1914). On the determination of size of family and of the distribution of characters in order of birth from samples taken through members of the sibships. *J. Roy. Statist. Soc.* **77**, 179.

HOGBEN, L. (1952). Selective limitation of sibship size. *Brit. J. Soc. Med.* **6**, 188.

KENDALL, M. G. & STUART, A. (1958). *The Advanced Theory of Statistics*, Vol. 1, p. 225. London: Griffin.

LENZ, W. (1959). Der Einfluss des Alters der Eltern und der Geburtennummer auf angeborne pathologische Zustände beim Kind. II Spezieller Teil. *Acta Genet.* **9**, 249.

LOWE, C. R. & MCKEOWN, T. (1950). The sex ratio of human births related to maternal age. *Brit. J. Soc. Med.* **4**, 75.

MEYERS, R. J. (1941). Effects of the war on the sex ratio at birth. *Amer. Sociological Rev.* **12**, 40.

NOVITSKI, E. & SANDLER, L. (1956). The relationship between parental age, birth order and the secondary sex ratio in humans. *Ann. Hum. Genet., Lond.*, **21**, 123.

PENROSE, L. S. (1934). The relative aetiological importance of birth order and maternal age in mongolism. *Proc. Roy. Soc.* B, **115**, 431.

PENROSE, L. S. (1954). Mongolian idiocy (mongolism) and maternal age. *Ann. N.Y. Acad. Sci.* **57**, 494.

SMITH, A. & RECORD, R. G. (1955). Maternal age and birth rank in the aetiology of mongolism. *Brit. J. Prev. Soc. Med.* **9**, 51.