Archives of Sexual Behavior, Vol. 31, No. 1, February 2002, pp. 63–71 (© 2002)

How Many Gay Men Owe Their Sexual Orientation to Fraternal Birth Order?

James M. Cantor, Ph.D.,[1] Ray Blanchard, Ph.D.,[1,2,5] Andrew D. Paterson, M.B., Ch.B.,[2,3] and Anthony F. Bogaert, Ph.D.[4]

Received October 16, 2000; revision received March 11, 2001; accepted September 1, 2001

In men, sexual orientation correlates with the number of older brothers, each additional older brother increasing the odds of homosexuality by approximately 33%. However, this phenomenon, the fraternal birth order effect, accounts for the sexual orientation of only a proportion of gay men. To estimate the size of this proportion, we derived generalized forms of two epidemiological statistics, the *attributable fraction* and the *population attributable fraction*, which quantify the relationship between a condition and prior exposure to an agent that can cause it. In their common forms, these statistics are calculable only for 2 levels of exposure: exposed versus not-exposed. We developed a method applicable to agents with multiple levels of exposure—in this case, number of older brothers. This noniterative method, which requires the odds ratio from a prior logistic regression analysis, was then applied to a large contemporary sample of gay men. The results showed that roughly 1 gay man in 7 owes his sexual orientation to the fraternal birth order effect. They also showed that the effect of fraternal birth order would exceed all other causes of homosexuality in groups of gay men with 3 or more older brothers and would precisely equal all other causes in a theoretical group with 2.5 older brothers. Implications are suggested for the gay sib-pair linkage method of identifying genetic loci for homosexuality.

KEY WORDS: attributable fraction; attributable risk; birth order; homosexuality; H-Y antigen; logistic regression; sexual orientation; sib-pair linkage method.

Epidemiological studies have repeatedly shown that older brothers increase the probability of homosexuality in later-born males (Blanchard, 1997, 2001; Jones & Blanchard, 1998). Older sisters, in contrast, do not affect the sexual orientation of later-born males, and neither older brothers nor older sisters affect the sexual orientation of later-born females. Because females are essentially invisible to this process, we have called it the *fraternal birth order effect*.

[1]Centre for Addiction and Mental Health, Toronto, Ontario, Canada.
[2]Department of Psychiatry, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada.
[3]Department of Genetics, The Hospital for Sick Children, Toronto, Ontario, Canada.
[4]Departments of Psychology and Community Health Sciences, Brock University, St. Catharines, Ontario, Canada.
[5]To whom correspondence should be addressed at CAMH—Clarke Division, 250 College Street, Toronto, Ontario, Canada M5T 1R8; e-mail: ray_blanchard@camh.net.

The fraternal birth order effect has been demonstrated not only in ordinary homosexual community volunteers (Blanchard & Bogaert, 1996a,b; Blanchard, Zucker, Siegelman, Dickey, & Klassen, 1998; Ellis & Blanchard, 2001; Robinson & Manning, 2000; see also Blanchard & Bogaert, 1997a; Purcell, Blanchard, & Zucker, 2000; Williams et al., 2000), but also in atypical homosexual groups who differ as widely as possible in their own characteristics and in the characteristics of their desired partners. Older brothers increase the probability that male-to-female transsexuals will be sexually attracted to men rather than women (Green, 2000; see also Zucker et al., 1997), that pedophiles will be sexually attracted to boys rather than girls (Blanchard et al., 2000; Bogaert, Bezeau, Kuban, & Blanchard, 1997), and that sex offenders against adults and pubescents will offend against males rather than females (Blanchard & Bogaert, 1998).

The diversity of the samples in which this phenomenon has been demonstrated makes fraternal birth order

0004-0002/02/0200-0063/0 © 2002 Plenum Publishing Corporation

(or the underlying variable it reflects) the most widespread factor in homosexual development that has yet been identified. There are, however, other measures of the theoretical importance of this phenomenon, in particular, the proportion of homosexual men who acquired their sexual orientation from the fraternal birth order effect, as opposed to some other agent. In this paper, we derive the generalized forms of epidemiological statistics needed to estimate this proportion, and we apply the generalized approach to a large, matched sample of homosexual and heterosexual volunteers.

A statistical approach that is capable of answering one empirical question is obviously capable of answering others that follow the same general form and satisfy the same underlying assumptions. We therefore briefly discuss the application of our method to similar research problems. We also discuss the implications of our empirical findings for genetic studies of sexual orientation.

THE ATTRIBUTABLE FRACTION AND POPULATION ATTRIBUTABLE FRACTION

The *attributable fraction* and *population attributable fraction* statistics provide intuitive measures of the magnitude of the relationship between two dichotomous variables. As typically used in epidemiology, these two variables are disease state (present vs. absent) and exposure to a pathogen (exposed vs. never exposed). These statistics are applied to diseases that can be caused by more than one pathogen. They quantify the relationship between the disease and any one specific pathogen, which we will call the target pathogen.

These statistics may be explained as follows. The population of all persons with a given disease can be divided into three groups: those who were exposed to the target pathogen and got the disease because of the exposure (*causally exposed*), those who were exposed to the target pathogen but actually got the disease from some other pathogen (*coincidentally exposed*), and those who were never exposed to the target pathogen and therefore necessarily got the disease from some other pathogen (*never exposed*). The attributable fraction is the ratio of diseased people who got the disease from the pathogen to the number of diseased people who were exposed to the pathogen, *causally exposed*/(*causally exposed* + *coincidentally exposed*). The population attributable fraction is the ratio of people who got the disease from the pathogen to the total number of diseased people, *causally exposed*/(*causally exposed* + *coincidentally exposed* + *never exposed*). The population attributable fraction can be thought of as the percentage of cases of disease (potentially) preventable by a total elimination of exposure in the entire population (Gefeller, 1992).

The population attributable fraction may be equivalently expressed as the product of the attributable fraction and the prevalence of exposure among diseased cases. This form makes it clear that exposure to a potent agent may be strongly related to a disease, yet account for only a small proportion of the existing cases because of the rarity of exposure to it. The historical use of the term "population" here is unfortunate, because it suggests that "attributable fraction" refers to a sample characteristic, and "population attributable fraction," to the corresponding population parameter. As just shown, however, the population attributable fraction actually conveys a different type of information.

Although the attributable fraction and population attributable fraction are most frequently applied in epidemiological analyses of diseases and their causes, these statistics may just as readily be applied to variables that do not describe pathology. Therefore, neutral terms will be used for interpretation here: *agent* rather than pathogen to describe the independent variable (IV), and *condition* rather than disease to describe the dependent variable (DV). For the same reason, we will refer to the statistics as attributable fractions rather than as *attributable risks*, or the many other names that have been used (Greenland & Robbins, 1988; Walter, 1978).

The case of two dichotomous variables is usually depicted with a 2×2 table (Fig. 1). The odds ratio in this case may then be expressed as

$$\frac{ad}{bc},$$

and the attributable fraction may be expressed as

$$\frac{\text{prevalence}_{\text{exposed}} - \text{prevalence}_{\text{unexposed}}}{\text{prevalence}_{\text{exposed}}} \quad \text{or} \quad \frac{\frac{d}{c+d} - \frac{b}{a+b}}{\frac{d}{c+d}}.$$

As noted earlier, the population attributable fraction is the simple product of the attributable fraction and the

	Condition Not Present	Condition Present
Not Exposed	*a*	*b*
Exposed	*c*	*d*

Fig. 1. Organization of data for a dichotomous condition and dichotomous exposure status. Cell values represent number of cases. Prevalence of the condition among cases without any exposure (i.e., baseline prevalence), $p_0 = b/(a + b)$. Prevalence of the condition among cases experiencing exposure to the agent, $p_1 = d/(c + d)$.

prevalence of exposure to the agent among the cases with the condition, or

$$\left(\frac{\frac{d}{c+d} - \frac{b}{a+b}}{\frac{d}{c+d}}\right)\left(\frac{d}{b+d}\right).$$

Unfortunately, these calculations of the attributable fractions are limited to IVs that possess exactly two levels. That is, data regarding the exposure status of a case must be dichotomized into exposed versus not-exposed. Although dichotomies appropriately capture information regarding many factors, they require researchers to collapse multilevel data into less precise, all-or-none categories.

Generalized forms of the attributable fraction that can accommodate multilevel variables are required to utilize the information inherent in multilevel data. Park (1981) suggested a method for quantifying the relation between a dichotomous IV and a multilevel DV. Park's method does not, however, apply to multilevel IVs. The problem of multiple-level IVs (i.e., multiple exposures) was discussed by Denman and Schlesselman (1983). Their suggested solution employed individual odds ratios, with a separate ratio estimated for each level of the IV relative to the unexposed group. Denman and Schlesselman's approach does not require a linear increase in odds between each adjacent level of the IV (i.e., a constant odds ratio for the sample). When, however, the relation between a condition and number of exposures is linear in the odds—as appears to be the case for sexual orientation and number of older brothers—that approach may not be the most efficient or accurate.

We therefore undertook to derive an attributable fraction for multiple exposures, a statistic that would use the odds ratio from a prior logistic regression analysis to directly calculate the attributable fraction for any given number of exposures. In this context, *exposures* refer broadly to occurrences of the causal agent, whether they impinge on the case directly or affect the case via their cumulative effect on some intervening variable. The latter situation is frequently encountered in birth order research, where the probability of a fetus developing certain diseases increases with the number of prior fetuses to which the maternal uterus is exposed. Children of later pregnancies are more likely to develop macrosomia (e.g., Babinszki et al., 1999), mental retardation (e.g., Flannery & Liederman, 1994), Down's syndrome (e.g., Schimmel, Eidelman, Zadka, Kornbluth, & Hammerman, 1997), and diabetes (e.g., Tuomilehto, Podar, Tuomilehto-Wolf, & Virtala, 1995). In these examples, prior fetuses would not affect the subsequent fetus—that is, the "case"—directly, but rather through the intervening variable of cumulative changes in the uterine environment.

DERIVATION OF THE ATTRIBUTABLE FRACTION FOR MULTIPLE EXPOSURES

Where the attributable fraction for two dichotomous variables may be depicted in a 2×2 table, situations in which cases may experience any number of exposures may be depicted in a table with $2 \times (N + 1)$ cells, where N indicates the maximum number of times any case has been exposed to the agent (Fig. 2). Let S be the odds ratio for the overall set of exposure levels, with the assumption of a constant increase in the odds of the condition being present between each level of the IV. Then, let odds_n be the odds of the condition being present after n exposures, with odds_0 being the odds of the condition being present with no exposure to the agent. Note that odds_0 may also be thought of as the baseline odds of a case having the condition or as the odds of having the condition because of a factor other than the agent represented by the IV.

After a single exposure to an agent, the odds of developing a condition increase by a factor of S, that is,

$$\text{odds}_1 = \text{odds}_0 \cdot S.$$

After a second exposure to an agent, the odds of developing the condition increase by a factor of S once again:

$$\text{odds}_2 = \text{odds}_0 \cdot S \cdot S.$$

	Condition Not Present	Condition Present
0 exposures	a_0	b_0
1 exposure	a_1	b_1
2 exposures	a_2	b_2
⋮	…	…
N exposures	a_N	b_N

Fig. 2. Organization of data for a dichotomous condition and an arbitrary number of exposures. Cell values represent number of cases. Prevalence of the condition among cases without any exposure, $p_0 = b_0/(a_0 + b_0)$. Prevalence of the condition among cases experiencing n exposures to the agent, $p_n = b_n/(a_n + b_n)$.

In general, the odds of the condition being present after n exposures is expressible as

$$\text{odds}_n = \text{odds}_0 \cdot S^n. \quad (1)$$

However, the attributable fraction, whether for single or for multiple exposures, is expressed in terms of probabilities or prevalence rates, rather than odds. Equation (1) may be recast in terms of probabilities by noting that odds = probability/(1 − probability). Replacing odds_n and odds_0 respectively with p_n and p_0, where p_n is the probability of the condition being present after n exposures and p_0 is the baseline probability of the condition, yields, with simplification,

$$p_n = \frac{1}{1 + \left(\frac{1}{p_0} - 1\right)/S^n}. \quad (2)$$

As noted earlier, the attributable fraction is the difference in prevalence (or probability) of cases between the exposed and nonexposed groups, expressed as a proportion of the prevalence among the exposed. This calculation is true in general. The attributable fraction for any individual level of the IV, G_n, is the difference in prevalence of cases between that level and the baseline level, expressed as a proportion of the prevalence at that level of exposure:

$$G_n = \frac{p_n - p_0}{p_n} = 1 - \frac{p_0}{p_n}.$$

G_n may be expressed in terms of S and p_0 using Eq. (2) to substitute for p_n, providing an equation for the attributable fraction for any n:

$$G_n = 1 - \frac{p_0}{1/\left[1 + \left(\frac{1}{p_0} - 1\right)/S^n\right]}$$

$$G_n = (1 - p_0)\left(1 - \frac{1}{S^n}\right). \quad (3)$$

Thus, each value of G_n can be calculated directly with Eq. (3), from only the baseline prevalence and overall odds ratio.

The population attributable fraction can then be calculated as the weighted average of each G_n, with weights assigned by the distribution of exposure levels in the population. If D_n is the proportion of cases that have experienced n exposures to the agent, then the population attributable fraction is

$$\sum_{n=0}^{n=N} G_n \cdot D_n, \quad (4)$$

or, in vector notation, the cross product of the column vectors **G** and **D**, **D′G**.

THE PROPORTION OF "OLDER-BROTHER-TYPE" GAYS

The proportion of gay men who owe their sexual orientation to the fraternal birth order effect was estimated by applying the foregoing method to unpublished data from Blanchard and Bogaert (1996b). The participants in this study were 302 homosexual men who were individually matched on year of birth with 302 heterosexual men. All participants described their race as White and reported that they were single births. None was adopted, had any maternal half-siblings, or expressed any doubt that he knew of all children born to his mother. Table I presents this sample broken down by sexual orientation and by number of older brothers.

Equation (3) requires two values, the baseline prevalence of the condition under investigation and the odds ratio for its increase between each level of exposure to the agent. The overall prevalence of homosexuality in the adult male population is probably somewhere between 2 and 3% (e.g., ACSF Investigators, 1992; Billy, Tanfer, Grady, & Klepinger, 1993; Fay, Turner, Klassen, & Gagnon, 1989; Johnson, Wadsworth, Wellings, Bradshaw, & Field, 1992; Laumann, Gagnon, Michael, & Michaels, 1994), and so we have estimated the baseline prevalence (or the prevalence of homosexuality among men with 0 older brothers) to be 2%. Logistic regression analysis of Table I data reveals an odds ratio of 1.33. Substituting these values into Eq. (3) produces the attributable fraction values, G_n, listed in Table II. Thus, for example, $G_2 \cong 0.43$; that is, 43% of the homosexual men with two older brothers in this sample can attribute their homosexuality to the older brother effect.

The frequency distribution of older brothers among homosexual men is estimated by the distribution in the sample and is expressed in Table II as a proportion of the total size of the sample of homosexual participants. The sum of the products of these pairs of elements yields a population attributable fraction of 15.1%. Thus, about

Table I. Sexual Orientation and Number of Older Brothers[a]

	Sexual orientation	
Older brothers	Heterosexual	Homosexual
0	198	165
1	75	86
2	19	34
3	8	12
4	1	4
5+	1	1

[a] From Blanchard and Bogaert (1996b).

Table II. Attributable Fraction of Homosexuality for Older Brothers and the Frequency Distribution of Older Brothers in the Homosexual Group

No. of older brothers, n	Attributable fraction, G_n	Frequency distribution, D_n
0	.000	.546
1	.243	.285
2	.426	.113
3	.563	.040
4	.667	.013
5+	.745	.003

one out of every seven homosexual men in this sample can attribute his homosexuality to the older brother effect.

Additional calculations showed that variation in the baseline prevalence—at least within the range of plausible values—makes very little difference in the results. Re-estimating the percentage of older-brother-type gays using a baseline prevalence of 1% yields a figure of 15.2%; re-estimating with a baseline prevalence of 4% yields a figure of 14.8%.

The precise form of the function relating the odds of homosexuality to a proband's number of older brothers is another issue that should be examined. Our estimate of a 33% increase in the odds per older brother was based on the assumption that this function is linear. This assumption is justified by the results of the following reanalysis of raw data from Blanchard and Bogaert (1996b). This reanalysis involved rerunning the logistic regression analysis from that study. As in the original study, the criterion variable was sexual orientation, coded dichotomously as heterosexual or homosexual. In the reanalysis, the predictor variables were the participant's number of older brothers, the squared number of older brothers, and the cubed number of older brothers (i.e., the linear, quadratic, and cubic terms for older brothers, respectively). With the linear term already entered into the regression equation, addition of the quadratic term produced no significant improvement, $\chi^2(1) = 0.24$, $p = .62$; and with the linear and quadratic terms already entered into the equation, addition of the cubic term produced no significant improvement, $\chi^2(1) = 0.23$, $p = .63$. Very similar results were obtained with a reanalysis of raw data from Blanchard et al. (1998). It therefore appears that a linear function best describes the relation over the range of values observed.

It seems likely that in a sample including a substantial proportion of participants with many older brothers, the function would prove curvilinear; that is, after 8 or 10 older brothers, additional older brothers would produce less or no further increase in the odds of homosexuality. Such a sample would not, however, be expected from a modern, industrialized population, either now or in the imaginable future. It is therefore reasonable to model the relation as a linear one, at least for contemporary samples.

THE AF_{50}

The development of Eq. (3), a continuous function for the attributable fraction at any n, makes possible another useful metric by which to describe the association between a condition and an agent associated with it. Because baseline frequency (p_0) is a positive constant, and because the contribution of exposures to the agent (G_n) is zero at $n = 0$ and increases as n increases, there will eventually be a point at which the fraction attributable to the exposures to the agent equals (and then exceeds) the fraction attributable to all the other (baseline) factors. That is, at some n, $p_n = 2p_0$. For cases at this n, the agent has contributed as much to the prevalence of the condition as did all the other effects that were present at baseline (i.e., at $n = 0$). We term this point the AF_{50}, because 50% of the cases are attributable to the agent. The AF_{50} is located by setting Eq. (3) to $2p_0$ and solving for n, that is,

$$(1 - p_0)\left(1 - \frac{1}{S^n}\right) = 2p_0$$

and thus,

$$n = \frac{\log(2 - 2p_0) - \log(1 - 2p_0)}{\log S}. \qquad (5)$$

Using the previous example estimating the baseline prevalence of homosexuality as 2% and the odds ratio of 1.33 from Table I, substitution into Eq. (5) yields an n of 2.503 older brothers. That is, among gay men with more than 2.5 older brothers, sexual orientation is more attributable to the fraternal birth order effect than to all other possible effects combined.

DISCUSSION

Our main finding is that roughly one gay man in seven owes his sexual orientation to the fraternal birth order effect. This shows that the contribution of fraternal birth order to the sum total of gay men is more than negligible.

Alternative strategies for quantifying the magnitude of the fraternal birth order effect in intuitively comprehensible terms treated number of older brothers as if it were a continuous variable, that is, as if someone could have fractions of an older brother. These analyses showed that a boy with 2.5 older brothers would be twice as likely to be gay as a boy with 0 older brothers and that, for mathematically related reasons, half of all gay men with 2.5 older brothers would not have been gay if they had

had 0 older brothers (i.e., the attributable fraction equals 50% when the number of older brothers is 2.5). This also means that among men with 2.5 older brothers, the fraternal birth order effect would equal all other causes of homosexuality combined. Previous work along these lines, which used informal methods and which happened to produce values that closely approximated integers, showed that a boy with 4 older brothers would be three times as likely to be gay as a boy with 0 older brothers (Blanchard, 2001).

The foregoing conclusions rest on the assumption that older brothers cause homosexuality, whether directly or indirectly. This assumption seems to us justified on a variety of logical and empirical grounds. From a purely mathematical standpoint, the observed correlation between homosexuality and older brothers could arise in three basically different ways: the former could cause the latter; the latter could cause the former; or both could be caused by some third variable. We can eliminate the first possibility on logical grounds—a man's homosexuality cannot operate backwards in time to give him additional older brothers. That leaves only one competing interpretation, namely, that homosexuality and older brothers are correlated only because both are caused by some third variable. The question then arises: What third variable?

One possibility might seem, at least at first glance, to be parental age. A man's fraternal birth order naturally correlates substantially with the age of his parents at the time of his birth. This raises the possibility that the seeming association of fraternal birth order and sexual orientation is merely a statistical artifact arising from the correlation of both with parental age, and that the important connection is between parental age and sexual orientation. A genetic explanation along these lines was suggested by Raschka (1995), who argued that a higher paternal age might reflect an increased mutation rate in the spermatogenesis of older fathers. The first problem with positing parental age as the hypothetical third variable is that several empirical studies have shown directly that the relation between number of older brothers and male homosexuality is not an artifact of higher maternal or paternal age (Blanchard & Bogaert, 1996a,b, 1997b, 1998; Blanchard & Sheridan, 1992; Bogaert et al., 1997). The second problem with positing parental age as the hypothetical third variable is logical (Jones & Blanchard, 1998). Parents' ages at the birth of a boy and that boy's birth order among his siblings are, as already noted, strongly correlated. The correlation, however, is essentially the same for both sexes. Hence, if homosexuality is directly related to advanced maternal or paternal age, gay men should tend to be born late with respect to both their brothers and their sisters. They are not, however; they are born late only with respect to their brothers. It therefore appears that parental age cannot explain the relation between fraternal birth order and sexual orientation.

The search for a third variable must therefore turn to other possibilities. These are rather difficult to envision, even if one permits the positing of hitherto unobserved phenomena. Imagine, for example, there exists some condition that causes a man to produce extraordinarily successful Y-bearing sperm. Such men sire large numbers of sons, both in relation to other men and in relation to their own number of daughters. The condition also, however, predisposes them to sire homosexual sons. The gay sons of such fathers would, in fact, have an excess of older brothers. Thus, this imaginary condition (a hypothetical third variable) does seem capable of accounting for the data until one reflects that such gay sons would have an equally large excess of younger brothers. It could not, therefore, explain the crucial finding that gay men have an excess only of older brothers. It seems to us, in summary, that alternative explanations of the correlation between fraternal birth order and homosexuality are either so clearly incorrect or else so difficult to conceive that the most plausible interpretation is the simple one that older brothers make some causal contribution to homosexuality in later-born males.

Dissimilar theories of the fraternal birth order effect converge on the conclusion that our estimates of its magnitude are likely to be on the conservative side. The reason for this, which we discuss below, is that these theories imply that the variable of real interest is not a man's number of live-born older brothers but something slightly different, for which his number of live-born older brothers is an imperfect measure. One theory implies that the variable of real interest is the number of prior male fetuses carried by a man's mother, whether these resulted in a live birth or not. A second theory implies that the variable of real interest is the number of older brothers that were actually in a man's environment when he was growing up, not his total number of older brothers.

Blanchard and Bogaert (1996b) hypothesized that the correlation of fraternal birth order with sexual orientation in males reflects the progressive immunization of some mothers to Y-linked minor histocompatibility antigens (H-Y antigens) by each succeeding male fetus, and the concomitantly increasing effects of anti-H-Y antibodies on the sexual differentiation of the brain in each succeeding male fetus (Blanchard & Bogaert, 1996b). This hypothesis rests partly on the argument that a woman's immune system would appear to be the biological system most capable of "remembering" the number of male (but not female) fetuses that she has previously carried and of progressively altering its response to the next fetus

according to the current tally of preceding males. It also rests on the finding that sexual orientation in females, who do not express H-Y antigens and would therefore not be targets of anti-H-Y antibodies in utero, is unrelated to their number of older brothers. Various lines of animal experimental evidence and human clinical evidence that bear on the plausibility of this hypothesis have been summarized by Blanchard and Klassen (1997). Recently, Blanchard and Bogaert's basic assumption that the fraternal birth order effect operates in the prenatal environment has been bolstered by the finding that homosexual males with older brothers weigh substantially less at birth than heterosexual males with older brothers (Blanchard, 2001; Blanchard & Ellis, 2001).

If the maternal immune hypothesis is correct, and if—as animal evidence suggests (Epstein, Smith, & Travis, 1980; Krco & Goldberg, 1976; Shelton & Goldberg, 1984; White, Anderson, & BonDurant, 1987; White, Lindner, Anderson, & BonDurant, 1983)—male fetuses begin expressing H-Y antigens very early in development, then spontaneously aborted male fetuses might also immunize the mother and thereby augment the probability of homosexuality in her subsequent male offspring. That would mean that a calculation based on the observed relation between a man's sexual orientation and his number of (live-born) older brothers underestimates the proportion of gays who owe their orientation to the fraternal birth order effect, or more precisely, to the mechanism underlying that effect. On that basis, one might argue that the true percentage of older-brother-type gays is more likely to lie above than below our present estimate.

A very different class of explanations assumes that the fraternal birth order effect operates postnatally, in the environment of rearing. The most popular example of these is the hypothesis that sexual interaction with older males increases a boy's probability of developing a homosexual orientation, and that a boy's chances of engaging in such interactions increase in proportion to his number of older brothers (Jones & Blanchard, 1998). If this hypothesis is correct, then one should only count, as older brothers, prior-born brothers who had the opportunity to interact sexually with the subject. Older brothers who left the family home or died before the subject was born, or while the subject was still an infant, represent statistical error that would reduce the apparent correlation between fraternal birth order and sexual orientation. There is no way to reliably identify and exclude such brothers in the data of Blanchard and Bogaert (1996b) or in any other data set known to us, however. Therefore this explanation, like most other psychosocial explanations, also implies that the available data underestimate the true percentage of older-brother-type gays.

As indicated in the beginning of the paper, our findings have implications for the gay sib-pair linkage method of identifying genetic loci for homosexuality (e.g., Hamer, Hu, Magnuson, Hu, & Pattatucci, 1993; Hu et al., 1995; Rice, Anderson, Risch, & Ebers, 1999). The present study shows that older-brother-type gays (in genetic terminology, *phenocopies*) add a considerable amount of noise to gay sib-pair analyses, and they add it nonrandomly. In the best-case scenario (best case for the geneticist), a gay sib-pair would consist of the first two boys born in a family. Table II shows, however, that the younger of the pair already has a 24% chance of being gay for nongenetic reasons. If a gay sib-pair consists of the second and third boys born in a family, the elder has a 24% chance of being gay for nongenetic reasons, and the younger has a 43% chance of being gay for nongenetic reasons. This analysis assumes, of course, that the fraternal birth order and familiality effects are independent; there are little data that bear on this assumption, but what there are suggest that the effects are independent (Blanchard & Bogaert, 1997a). It therefore appears that the fraternal birth order phenomenon may significantly increase the difficulty of finding genetic linkage using a gay sib-pair design.

The foregoing problem might be approached in parametric linkage analyses of gay sib-pair data by specifying different liability classes (specifically, phenocopy rates) for different individuals. This is usually used to provide age- and sex-dependent penetrances, but it could also be used to assign different phenocopy rates to sibs with different birth orders (Ott, 1991). An analogous adjustment might be made in so-called nonparametric analyses of gay sib-pair data, by incorporating covariates or using a weighting factor based on the birth orders of the individual sibs (Dawson, Kaplan, & Elston, 1990; Flanders & Khoury, 1991; Greenwood & Bull, 1999; Yang & Khoury, 1997).

Three assumptions underlying the attributable fraction for multiple exposures and the AF_{50} should be made explicit, for the benefit of researchers who might consider applying these statistics to other problems. First is the assumption of a constant odds ratio (i.e., a linear increase in odds) over increasing levels of the IV. As we have shown, this assumption can readily be tested by examining the statistical significance of the higher order terms for the IV in the logistic regression equation. Second is the assumption that the odds of the condition occurring increase rather than decrease with each exposure to the agent, that is, $S > 1$. Agents that are associated with decreased odds of developing the condition, that is, $S < 1$, are referred to as *protective factors*; the statistics developed here cannot be meaningfully interpreted for instances of

protective factors. Third is the assumption that the case of zero exposures, p_0, provides a meaningful baseline prevalence of the condition.

ACKNOWLEDGMENTS

The authors thank J. Michael Bailey, Scott Hershberger, and Edward Miller for their comments on earlier drafts of this paper. This research was supported by Social Sciences and Humanities Research Council of Canada Grant 410-99-0019 to Ray Blanchard and by a postdoctoral fellowship award from the CAMH Foundation and the Ontario Ministry of Health to James M. Cantor.

REFERENCES

ACSF Investigators. (1992). AIDS and sexual behaviour in France. *Nature, 360,* 407–409.

Babinszki, A., Kerenyi, T., Torok, O., Grazi, V., Lapinski, R. H., & Berkowitz, R. L. (1999). Perinatal outcome in grand and great-grand multiparity: Effects of parity on obstetric risk factors. *American Journal of Obstetrics and Gynecology, 181,* 669–674.

Billy, J. O. G., Tanfer, K., Grady, W. R., & Klepinger, D. H. (1993). The sexual behavior of men in the United States. *Family Planning Perspectives, 25,* 52–60.

Blanchard, R. (1997). Birth order and sibling sex ratio in homosexual versus heterosexual males and females. *Annual Review of Sex Research, 8,* 27–67.

Blanchard, R. (2001). Fraternal birth order and the maternal immune hypothesis of male homosexuality. *Hormones and Behavior, 40,* 105–114.

Blanchard, R., Barbaree, H. E., Bogaert, A. F., Dickey, R., Klassen, P., Kuban, M. E., et al. (2000). Fraternal birth order and sexual orientation in pedophiles. *Archives of Sexual Behavior, 29,* 463–478.

Blanchard, R., & Bogaert, A. F. (1996a). Biodemographic comparisons of homosexual and heterosexual men in the Kinsey interview data. *Archives of Sexual Behavior, 25,* 551–579.

Blanchard, R., & Bogaert, A. F. (1996b). Homosexuality in men and number of older brothers. *American Journal of Psychiatry, 153,* 27–31.

Blanchard, R., & Bogaert, A. F. (1997a). Additive effects of older brothers and homosexual brothers in the prediction of marriage and cohabitation. *Behavior Genetics, 27,* 45–54.

Blanchard, R., & Bogaert, A. F. (1997b). Drs. Blanchard and Bogaert reply [Letter to the editor]. *American Journal of Psychiatry, 154,* 137.

Blanchard, R., & Bogaert, A. F. (1998). Birth order in homosexual versus heterosexual sex offenders against children, pubescents, and adults. *Archives of Sexual Behavior, 27,* 595–603.

Blanchard, R., & Ellis, L. (2001). Birth weight, sexual orientation, and the sex of preceding siblings. *Journal of Biosocial Science, 33,* 451–467.

Blanchard, R., & Klassen, P. (1997). H-Y antigen and homosexuality in men. *Journal of Theoretical Biology, 185,* 373–378.

Blanchard, R., & Sheridan, P. M. (1992). Sibship size, sibling sex ratio, birth order, and parental age in homosexual and nonhomosexual gender dysphorics. *Journal of Nervous and Mental Disease, 180,* 40–47.

Blanchard, R., Zucker, K. J., Siegelman, M., Dickey, R., & Klassen, P. (1998). The relation of birth order to sexual orientation in men and women. *Journal of Biosocial Science, 30,* 511–519.

Bogaert, A. F., Bezeau, S., Kuban, M., & Blanchard, R. (1997). Pedophilia, sexual orientation, and birth order. *Journal of Abnormal Psychology, 106,* 331–335.

Dawson, D. V., Kaplan, E. B., & Elston, R. C. (1990). Extensions to sib-pair linkage tests applicable to disorders characterized by delayed onset. *Genetic Epidemiology, 7,* 453–466.

Denman, D. W., & Schlesselman, J. J. (1983). Interval estimation of the attributable risk for multiple exposure levels in case-control studies. *Biometrics, 39,* 185–192.

Ellis, L., & Blanchard, R. (2001). Birth order, sibling sex ratio, and maternal miscarriages in homosexual and heterosexual men and women. *Personality and Individual Differences, 30,* 543–552.

Epstein, C. J., Smith, S., & Travis, B. (1980). Expression of H-Y antigen on preimplantation mouse embryos. *Tissue Antigens, 15,* 63–67.

Fay, R. E., Turner, C. F., Klassen, A. D., & Gagnon, J. H. (1989). Prevalence and patterns of same-gender sexual contact among men. *Science, 243,* 338–348.

Flanders, W. D., & Khoury, M. J. (1991). Extensions to methods of sib-pair linkage analyses. *Genetic Epidemiology, 8,* 399–408.

Flannery, K. A., & Liederman, J. (1994). A test of the immunoreactive theory of the origin of neurodevelopmental disorders: Is there an antecedent brother effect? *Developmental Neuropsychology, 10,* 481–492.

Gefeller, O. (1992). An annotated bibliography on the attributable risk. *Biometrics Journal, 34,* 1007–1012.

Green, R. (2000). Birth order and ratio of brothers to sisters in transsexuals. *Psychological Medicine, 30,* 789–795.

Greenland, G., & Robbins, J. M. (1988). Conceptual problems in the definition and interpretation of attributable fractions. *American Journal of Epidemiology, 128,* 1185–1197.

Greenwood, C. M., & Bull, S. B. (1999). Analysis of affected sib pairs, with covariates—with and without constraints. *American Journal of Human Genetics, 64,* 871–885.

Hamer, D. H., Hu, S., Magnuson, V. L., Hu, N., & Pattatucci, A. M. L. (1993). A linkage between DNA markers on the X-chromosome and male sexual orientation. *Science, 261,* 321–327.

Hu, S., Pattatucci, A. M. L., Patterson, C., Li, L., Fulker, D. W., Cherny, S. S., et al. (1995). Linkage between sexual orientation and chromosome Xq28 in males but not in females. *Nature Genetics, 11,* 248–256.

Johnson, A. M., Wadsworth, J., Wellings, K., Bradshaw, S., & Field, J. (1992). Sexual lifestyles and HIV risk. *Nature, 360,* 410–412.

Jones, M. B., & Blanchard, R. (1998). Birth order and male homosexuality: An extension of Slater's Index. *Human Biology, 70,* 775–787.

Krco, C. J., & Goldberg, E. H. (1976). H-Y (male) antigen: Detection on eight-cell mouse embryos. *Science, 193,* 1134–1135.

Laumann, E. O., Gagnon, J. H., Michael, R. T., & Michaels, S. (1994). *The social organization of sexuality: Sexual practices in the United States.* Chicago: University of Chicago Press.

Ott, J. (1991). *Analysis of human genetic linkage* (2nd ed.). Baltimore, MD: Johns Hopkins University Press.

Park, C. B. (1981). Attributable risk for recurrent events: An extension of Levin's measure. *American Journal of Epidemiology, 113,* 491–493.

Purcell, D. W., Blanchard, R., & Zucker, K. J. (2000). Birth order in a contemporary sample of gay men. *Archives of Sexual Behavior, 29,* 349–356.

Raschka, L. B. (1995). On older fathers [Letter to the editor]. *American Journal of Psychiatry, 152,* 1404.

Rice, G., Anderson, C., Risch, N., & Ebers, G. (1999). Male homosexuality: Absence of linkage to microsatellite markers at Xq28. *Science, 284,* 665–667.

Robinson, S. J., & Manning, J. T. (2000). The ratio of 2nd to 4th digit length and male homosexuality. *Evolution and Human Behavior, 21,* 333–345.

Schimmel, M. S., Eidelman, A. I., Zadka, P., Kornbluth, E., & Hammerman, C. (1997). Increased parity and risk of trisomy 21: Review of 37 100 live births. *British Medical Journal, 314,* 720–721.

Shelton, J. A., & Goldberg, E. H. (1984). Male-restricted expression of H-Y antigen on preimplantation mouse embryos. *Transplantation*, *37*, 7–8.

Tuomilehto, J., Podar, T., Tuomilehto-Wolf, E., & Virtala, E. (1995). Evidence for importance of gender and birth cohort for risk of IDDM in offspring of IDDM parents. *Diabetologia*, *38*, 975–982.

Walter, S. D. (1978). Calculation of attributable risk from epidemiological data. *International Journal of Epidemiology*, *7*, 175–182.

White, K. L., Anderson, G. B., & BonDurant, R. H. (1987). Expression of a male-specific factor on various stages of preimplantation bovine embryos. *Biology of Reproduction*, *37*, 867–873.

White, K. L., Lindner, G. M., Anderson, G. B., & BonDurant, R. H. (1983). Cytolytic and fluorescent detection of H-Y antigen on preimplantation mouse embryos. *Theriogenology*, *19*, 701–705.

Williams, T. J., Pepitone, M. E., Christensen, S. E., Cooke, B. M., Huberman, A. D., Breedlove, N. J., et al. (2000). Finger-length ratios and sexual orientation. *Nature*, *404*, 455–456.

Yang, Q., & Khoury, M. J. (1997). Evolving methods in genetic epidemiology. III: Gene-environment interaction in epidemiologic research. *Epidemiologic Reviews*, *19*, 33–43.

Zucker, K. J., Green, R., Coates, S., Zuger, B., Cohen-Kettenis, P. T., Zecca, G. M., et al. (1997). Sibling sex ratio of boys with gender identity disorder. *Journal of Child Psychology and Psychiatry*, *38*, 543–551.