Articles

Prevalence and Patterns of Same-Gender Sexual Contact Among Men

ROBERT E. FAY, CHARLES F. TURNER,* ALBERT D. KLASSEN, JOHN H. GAGNON

The prevalence and patterns of same-gender sexual contact among men are key components of models of the spread of HIV infection and AIDS in the U.S. population. Previous estimates by Kinsey *et al.* from data collected between 1938 and 1948 have been widely criticized for inadequacies of sample design. New lower-bound estimates of prevalence developed from data from a national sample survey conducted in 1970 indicate that minimums of 20.3 percent of adult men in the United States in 1970 had sexual contact to orgasm with another man at some time in life; 6.7 percent had such contact after age 19; and between 1.6 and 2.0 percent had such contact within the previous year. Although these estimates incorporate adjustments for missing data, the likelihood of underreporting suggests that these estimates might be lower bounds on the prevalence of same-gender sex among men. Two sets of alternative estimates are derived to assess the sensitivity of these estimates to the assumptions made in imputing values to missing data. Detailed estimates are presented by frequency of contact, age, education, and marital status; and supporting estimates are derived from a 1988 national survey. Data from both the 1970 and 1988 surveys indicate that never-married men are more likely than other men to have had same-gender sexual contacts within the last year. The 1970 survey also indicates, however, that approximately half the men estimated to have such contacts are found among the more numerous population of currently or previously married men.

WIDELY USED PROJECTIONS OF THE MAGNITUDE OF THE human immunodeficiency virus (HIV) epidemic in the United States depend crucially on estimates of the size of the population of men who have sex with men. For example, the Coolfont (*1*) estimate that 1.5 million people in the United States were infected by HIV as of 1986 made use of data collected by Kinsey *et al.* (*2*) between 1938 and 1948 to estimate that "more than 2.5 million (4 percent) of U.S. men between 16 and 55 years of age are exclusively homosexual throughout their lives; an estimated 5–10 million more will have some homosexual contact" (*1*). The estimates of population size were multiplied by the presumed prevalence of HIV infection in these groups in order to derive an estimate of the number of men infected with HIV through same-gender sexual contacts. These same estimates have also been used in the more recent reviews of the prevalence of HIV infection (*3*).

Even 40 years ago, Kinsey's data were regarded as unsuitable for making such estimates (*4*). Kinsey gathered most of his cases by recruiting networks of friends through contact persons who offered him entree to institutional groups (for example, faculty members who introduced Kinsey to their classes) and through similar contacts that led to less institutionalized collections of persons (for example, Parent Teachers Associations). The sample was disproportionately drawn from the Midwest and from college campuses.

Since the Kinsey sample was not a probability sample, the data do not allow estimation of the characteristics of the national population with knowable margins of error. It is this point that is made most trenchantly in the major statistical reviews (*4*) of Kinsey's research.

The failure of the data of Kinsey *et al.* to satisfy our present needs does not, of course, detract from the pioneering efforts of these investigators to systematically study human sexual behavior among a broad collection of adult Americans. Their work has influenced the way in which we think about the varieties of human sexual behavior, and it hastened the adoption of systematic interview and research techniques in the study of human sexual behavior (*5*).

Nonetheless, the need for more contemporary data is evident (*6*). Below we describe an attempt to obtain an alternative estimate of the proportion of the U.S. male population who have had sexual experiences with another man sometime during their lifetime (specifically, sexual experiences with other males when either the respondent or his partner came to a sexual climax). These estimates of the frequency of same-gender sex are derived for selected age groups, educational levels, and marital statuses. Supporting estimates are derived from a national survey conducted in 1988 by Michael et al. (*7*), and the sensitivity of the 1970 estimates to the assumptions made in imputing values to missing data is explored.

The numerical estimates presented throughout this article reflect what respondents reported in response to a series of survey questions. Our estimates will provide an accurate reflection of the prevalence of actual behavior only if respondents answer the survey questions in an unbiased manner—that is to say, if they do not systematically distort their responses (although random errors of recall or reporting may occur). For many survey topics, this assumption is reasonable. Societal intolerance (*8*), however, may cause some survey respondents to conceal histories of same-gender sexual contact (*9*). Although there is no simple way of estimating the number of such respondents, it is likely that the number of men who conceal such experiences exceeds the number of respondents who (deliberately or through misunderstanding of survey instructions) report homosexual contacts that did not occur. Given this potential reporting bias, the numerical estimates presented in Tables 2 through 8 might be regarded as setting lower bounds—except for

R. Fay is a consultant and C. Turner is study director for the Committee on AIDS Research and the Behavioral, Social, and Statistical Sciences, National Research Council, National Academy of Sciences, Washington, DC 20418; A. Klassen is a research associate in the Department of Sociology, University of North Dakota, Grand Forks, ND 58202; and J. Gagnon is professor of sociology, SUNY at Stony Brook, Stony Brook, NY 11794.

*To whom correspondence should be addressed.

possible effects of sampling error—on the actual number of men who have sex with other men.

Survey Characteristics

Our estimates are derived by secondary analysis of survey data for a sample of 3018 U.S. adults interviewed in late 1970, and specifically for the 1450 men in this sample whose reported age was 21 years or older. These data were initially collected under the direction of the Kinsey Institute for Sex Research (of Indiana University). Only a brief and partial report of the original investigators' findings was published (*10, 11*). At present, these data appear to be the only extant source of detailed information on same-gender behaviors in a sample of the U.S. population (prior to the AIDS epidemic) derived from a sampling procedure that permits the calculation of meaningful error estimates (*12*).

In describing our analyses of these survey data, we purposefully use the term "estimates." Fifteen percent of the men in this sample did not report whether or not they had any same-gender experience, and an additional 6 percent (*13*) acknowledged some experience but did not respond completely to the relevant questions. Thus, any conclusions about prevalence or patterns of same-gender sexual behavior in the population require that one make inferences about behaviors among those men who did not respond completely to the survey questions.

In addition to trying to provide bases for making such inferences, we have further analyzed the sample of men with same-gender sexual experience by (i) frequency of contact, grouped into the categories "once, twice, rarely, occasionally, or fairly often" (categories into which respondents were asked to classify their contact in the survey); and by (ii) age at last contact, grouped into three categories: "less than 15 years old, 15 to 19, or 20 and older."

Obviously, in light of the current epidemic of HIV infections, these two variables fall short of the behavioral detail one might desire. Preceding awareness of the HIV virus by over a decade, the survey questionnaire was not designed to collect information on "safer sex" practices, such as the use of condoms, nor ask about the type of sexual contact (oral, anal, and so on). Nonetheless, the data may provide important information about the same-gender sexual experiences of the male population prior to the onset of the current epidemic.

Sample design and execution. According to documents provided by the National Opinion Research Center (NORC), the organization that did the original fieldwork, the survey design followed was that commonly employed for NORC's amalgam surveys (*14*). The universe for the survey sample was the noninstitutional population of the continental United States, age 21 and older. The sample was drawn using a multistage area probability design to the block or segment level. At the block or segment level, however, interviewers were given a prescribed travel pattern beginning at a random dwelling unit, and they were instructed to proceed in a specified direction until they had obtained completed interviews from a quota of persons in each of several categories. (Interviewers were restricted to a single interview per household.) Documentation for the 1970 survey indicates that interview quotas were established for men age 21 to 29, men 30 and over, employed women, and unemployed women. In a full-probability sample design, in contrast, the selection of respondents within blocks or segments would usually be accomplished by listing housing units (or eligible respondents) and sampling at random from this listing. An experimental comparison of estimates obtained by NORC's "probability sampling with quotas" versus full-probability sampling was carried out in 1975 and 1976. The results suggest that estimates derived from the two designs are quite similar (*15*).

We do not know what proportion of potential respondents refused to participate in the 1970 survey because this information was not assembled. Documents provided by NORC indicate that the survey's field staff perceived the refusal rates for two preliminary tests of this survey to be similar to refusal rates encountered in surveys of less sensitive topics (*16*).

To provide suggestive information on the adequacy of the sample design and execution, we compared the distributions of age, education, race, and marital status derived from the men in this sample to the distributions obtained in the 1970 census (Table 1). Although there is considerable agreement between the sample and the census, we note that there are some modest departures suggesting an overrepresentation of men 65 years of age and older (18.7 versus 14.5 percent), persons with 1 to 3 years of college education (14.5 versus 10.7 percent), and a more substantial overrepresentation of black men (16.1 percent in the Kinsey-NORC sample versus 9.0 percent in the census). The overrepresentation of blacks is thought to be caused, in part, by the use of the 1960 census as a sampling frame. Migration of white respondents to new housing units (not enumerated in the 1960 census) could have produced some underrepresentation of whites in the 1970 survey. The observed divergence, however, is sufficiently large to suggest that other factors may have operated (*17*).

Below we compare results obtained in the 1970 survey to comparable data on sexual behavior from the 1988 NORC General Social Survey, which employed a full-probability design.

Interview procedure. Respondents in the survey were interviewed in

Table 1. Age, race, education, and marital status distributions of the 1970 Kinsey-NORC sample and the 1970 census, excluding the population in group quarters. For age and race, comparisons are for the population aged 21 years and older. For education and marital status, the sample and the census are restricted to men 25 years of age and older (*31*).

Demographic characteristic	1970 census (%)	1970 Kinsey-NORC (%)
Age*		
21–24	9.3	9.1
25–29	11.4	9.5
30–44	29.7	31.2
45–64	35.2	31.5
≥65	14.5	18.7
Race*		
White and other	91.0	83.9
Black	9.0	16.1
Education†		
Less than high school	47.7	45.3
High school	27.9	26.6
College, 1 to 3 years	10.7	14.5
College, ≥4 years	13.6	13.5
Marital status†		
Married	83.0	82.7
Single	8.3	5.8
Divorced	3.3	3.4
Separated	1.7	2.2
Widowed	3.8	5.9

*The census tabulation excludes persons residing in group quarters to make it comparable to the NORC universe. The census estimate for ages 21 to 24 was computed by assuming that the proportion of males ages 21 to 24 who lived in group quarters was the same as the proportion in the age category 20 to 24 living in group quarters. (The relevant publication from the 1970 Bureau of Census, PC(2)-4E (*31*), tabulates persons residing in group quarters by age and gender using the 20–24 age group.) †Both the census and Kinsey-NORC samples have been restricted to men age ≥25. For the education distribution, all published census tabulations refer to this population. For marital status, required tabulations to remove persons in group quarters from the census do not allow precise identification of the population age ≥21. Since the marital status distribution is quite sensitive to age in this range, adjustments (as done for race) were not appropriate.

person at their residence. Interviewers were instructed (*18*) that "it is of the utmost importance that you be alone with the respondent during the interview session," and were given detailed procedures to obtain a private interview if a spouse or other person was present. In cases where it was "absolutely impossible to talk to the respondent in privacy," interviewers were advised that "it will be necessary to terminate the interview (graciously, of course)."

After a few opening questions about past participation in and opinions toward surveys of sexual attitudes and behavior, respondents were asked 12 questions requesting demographic information about themselves, their household, and their parents. These items were followed by approximately 70 questions largely concerned with attitudes toward various types of sexual behavior—for example, masturbation, premarital and extramarital intercourse, and same-gender sex. There were also some questions about other topics—for example, respondents were requested to rate their mother and father as "aggressive," "cool," and so forth. Social background questions were interspersed among these items. At the conclusion of this section of the interview, respondents were asked whether someone of the opposite sex had ever proposed having sexual relations with them, whether they had ever proposed sexual relations to someone of the opposite sex, and whether anyone of the same sex had ever "clearly proposed or attempted sexual relations" with them.

Subsequently, interviewers were instructed to give respondents a booklet of questions and to tell the respondents (*19*):

> In order to make a true evaluation of this entire survey, it is important to know something about people's experience. We would greatly appreciate your filling out this booklet, which is, of course, completely confidential. I will give you an envelope in which you, yourself, will seal the completed booklet. The answers from all sorts of people are really needed for statistical purposes.

The booklet's cover sheet included instructions reiterating the confidential nature of the survey. Except when the respondent requested help, the interviewers did not examine the booklet. The answers from the booklet were later matched to the previous part of the interview for analysis. The booklet contained a series of questions about the respondent's experience with (i) sex play as a child, (ii) masturbation, (iii) heterosexual activity before marriage, (iv) heterosexual experiences with spouse, (v) homosexual experiences, and (vi) sexual orientation and general enjoyment of sexual experiences.

The questions on homosexual experiences followed a sequence parallel to those of a previous series of questions on heterosexual experiences. The series of questions began with

> What was your age the first time you had a sexual experience with someone of the *same sex*, when either you or your partner came to a sexual climax?

The survey then continued with questions about the first partner's age, respondent's age at last experience, and frequency of experience, which was asked as follows:

> Was there a period of time when you had this experience fairly often, occasionally, or rarely, or did it happen only twice? (CIRCLE ONE.)

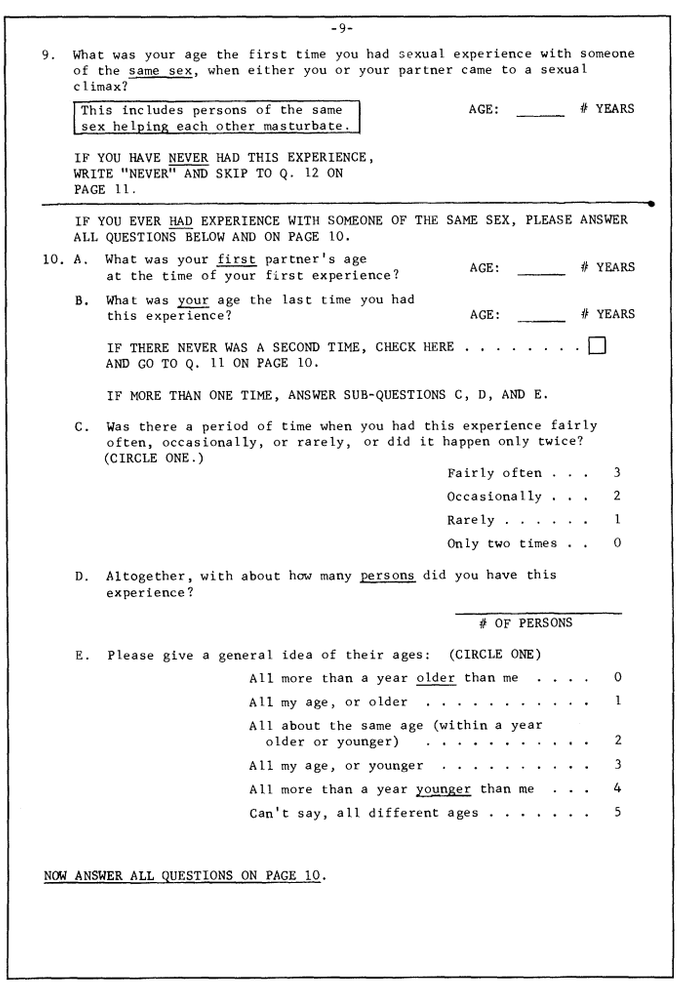
Other questions asked about the number of partners and their age relative to the respondent's age.

Because the questionnaire was filled out by the respondent and because it may have invited some misconceptions on the part of the respondents, we have reproduced the relevant page of the questionnaire (Fig. 1). (Questions on the following page of the questionnaire asked respondents about their feelings regarding their homosexual experiences and whether other persons knew about these experiences.) Clearly the questionnaire introduces some noteworthy possibilities for misinterpretation. For example, some respondents could have interpreted question 10B (Fig. 1) to refer to their last experience with their first partner, since the immediately preceding question (10A) asked for the age of the first partner.

Missing Data Analysis

It can be seen from Fig. 1 that respondents without any homosexual experience were directed in question 9 to enter "never" in the blank space labeled "age." However, of the 1,450 male respondents 21 years and older, a total of 223 skipped question 9 and all subsequent items on homosexual experience. In addition, 77 respondents indicated some homosexual experience by answering some of the associated questions but did not complete the box following 10B or answer question 10C to indicate the frequency of this experience. Among the 77 are a few cases for which a check in the box following 10B was treated as missing on account of inconsistencies with other responses.

Table 2 presents cross-classifications of reported values of age at last contact and frequency of contact, with our effort to interpret the missing responses. Part A shows the cross-classification of the two variables for 164 respondents reporting homosexual experience and providing valid answers to both questions. In a few instances, question 10B had not been answered directly, but an age at last experience of ≥20 may be inferred from the response to the age at first experience, question 9, of 20 years or more. The next four parts, B through E, present the results of our imputations for nonresponse



-9-

9. What was your age the first time you had sexual experience with someone of the same sex, when either you or your partner came to a sexual climax?

This includes persons of the same sex helping each other masturbate.

AGE: _____ # YEARS

IF YOU HAVE NEVER HAD THIS EXPERIENCE, WRITE "NEVER" AND SKIP TO Q. 12 ON PAGE 11.

IF YOU EVER HAD EXPERIENCE WITH SOMEONE OF THE SAME SEX, PLEASE ANSWER ALL QUESTIONS BELOW AND ON PAGE 10.

10. A. What was your first partner's age at the time of your first experience? AGE: _____ # YEARS

B. What was your age the last time you had this experience? AGE: _____ # YEARS

IF THERE NEVER WAS A SECOND TIME, CHECK HERE . . . . . ☐ AND GO TO Q. 11 ON PAGE 10.

IF MORE THAN ONE TIME, ANSWER SUB-QUESTIONS C, D, AND E.

C. Was there a period of time when you had this experience fairly often, occasionally, or rarely, or did it happen only twice? (CIRCLE ONE.)

Fairly often . . . 3
Occasionally . . . 2
Rarely 1
Only two times . . 0

D. Altogether, with about how many persons did you have this experience?

_____ # OF PERSONS

E. Please give a general idea of their ages: (CIRCLE ONE)

All more than a year older than me 0
All my age, or older 1
All about the same age (within a year older or younger) 2
All my age, or younger 3
All more than a year younger than me . . . 4
Can't say, all different ages 5

NOW ANSWER ALL QUESTIONS ON PAGE 10.

Fig. 1. Reproduction of page from self-administered questionnaire used in 1970 Kinsey Survey conducted by the National Opinion Research Center.

in different situations. The imputation took the form of an estimate of a probability distribution over possible values for the unreported variable or variables, rather than an assignment of an individual value as is often the practice in survey research generally. The assignment of probabilities has two advantages: (i) given our assumptions about the nature of the missing data, the assignment of probabilities gives a lower sampling variability than the assignment of single values; and (ii) this approach facilitates calculations, described below, to assess the joint effect of the sample design and imputation for missing values on the reliability of the estimates (*20*).

The first class of imputations (B in Table 2) were for 15 cases in which frequency of contact was reported but age at last contact was missing. For these cases, age at last contact was imputed according to the distributions of reported cases with similar age at first contact, (again grouped into the categories under 15, 15 to 19, and ≥20). When age at first contact was unavailable, the age at last contact was distributed according to complete cases with the same reported frequency of contact. Note that the distribution of imputed values is shifted somewhat away from the lowest age category compared to the distribution of completely reported cases.

The next class of imputations (C in Table 2) involved 30 cases in which age at last contact was reported but frequency of contact was missing. The distributions of frequency of homosexual contact were estimated from complete cases in the same grouping of age at last contact. Because most such incomplete cases represented instances in which the reported age at last contact was 20 or older, the imputed distribution for these cases is shifted toward the higher end of activity compared to the overall distribution for complete cases.

The 47 cases forming the next class (D in Table 2) contained some indication of homosexual experience, but responses were missing for both frequency of contact and age at last contact. The distribution of frequency of contact was first assumed for these cases from the overall distribution in parts A, B, and C. The imputation of age at last contact was similar to the first class of imputations. When available, age at first contact was again reflected in the imputations, or else frequency of contact was used. Note that the imputed joint probability distribution for age at last contact and frequency of contact closely reflects the association between these two variables observed for complete cases.

Finally, 223 cases offered no information on homosexual experience. In some of these cases, respondents had skipped large blocks of items on the self-administered questionnaire altogether, whereas in other cases respondents appeared to skip question 9 and move on to the questions following those on homosexual experience. To impute values for the homosexual experience variables in these cases, we made use of the association found between demographic characteristics, responses to attitude questions, and the reports of homosexual experience observed among respondents who completed these items. A total of 256 cases (that is, 164 + 15 + 30 + 47 in Table 2) of the other 1227 cases (1450 − 223), or 21 percent, had indicated some same-gender sexual experience. A logistic model was fit to these 1227 cases to predict the presence of any homosexual experience on the basis of demographic and attitude items collected by the interviewer during the previous part of the interview (*21*). The items measured attitudes toward homosexuality and laws against prostitution, and self-characterizations of approving attitudes about premarital and extra-marital sex compared to friends. A three-category classification—under 30, over 30 and never married, and over 30 and ever married—was also employed. These variables exhibited strong statistical associations with the reporting of same-gender sexual experience. When these relationships were applied to impute homosexual contact from the observed attitude and demographic data for the 223 nonrespondents, the total expected number of these persons with such contact was predicted to be 38.5 (17.3 percent), which is slightly below the rate observed in the rest of the sample. In fact, although response to the attitude items of a few of the nonrespondents paralleled that of respondents who reported homosexual experience, as a group, the nonrespondents expressed attitudes somewhat less typical of those who reported homosexual experience than the rest of the sample.

The probability of homosexual experience predicted by the logistic model was distributed over the 15 cells of the cross-classification of age at last contact and frequency according to the methods used on the other cases with missing data. This distribution (E of Table 2) closely mirrors that of complete cases.

It might be expected that respondents with no homosexual experience would find question 9 easy to answer compared to those with experience. Thus, it is somewhat surprising that the model predicts so few instances of same-gender sexual experience among the 223 nonrespondents. Several explanations are possible, however. Respondents skipping not only this item but large blocks of the self-administered questionnaire possibly may have done so for reasons other than homosexual experience. Secondly, respondents

Table 2. Homosexual experience by age at last contact and frequency of contact during the period of peak activity. Part A of this table provides counts (n) for cases with complete reports. Parts B through E show the counts imputed for cases with different types of nonresponse; fractional entries in these panels arise from the imputation procedure. Part F sums the reported and imputed case counts, and Part G presents these sums as a percentage of the total sample of men (≥21 years) in the sample. Note that each interior cell in (G) presents estimates as a percentage of the entire sample, not as a percentage of the row or column.

Age at last contact	Frequency of contact					Total
	(1) Once	(2) Twice	(3) Rare	(4) Occas.	(5) Fairly often	
	A, *Complete cases with no imputation* (n)					
<15	27	11	13	20	5	76
15–19	8	7	10	11	7	43
≥20	5	6	12	13	9	45
Total	40	24	35	44	21	164
	B, *Imputation of age at last contact only* (n)					
<15	1.4	0.5	0.4	1.3	1.0	4.4
15–19	0.4	0.8	1.3	1.4	1.3	5.2
≥20	0.2	0.8	1.4	1.3	1.7	5.4
Total	2.0	2.0	3.0	4.0	4.0	15.0
	C, *Imputation of frequency of activity only* (n)					
<15	2.1	0.9	1.0	1.6	0.4	6.0
15–19	0.4	0.3	0.5	0.5	0.3	2.0
≥20	2.4	2.9	5.9	6.4	4.4	22.0
Total	4.9	4.1	7.4	8.4	5.1	30.0
	D, *Imputation of both variables for those with some experience* (n)					
<15	5.6	2.6	3.4	4.9	1.7	18.2
15–19	3.0	2.2	3.2	3.8	2.3	14.5
≥20	2.4	2.0	3.4	4.0	2.5	14.3
Total	11.0	6.8	10.0	12.6	6.6	47.0
	E, *Imputation of both variables for 223 men with unknown experience* (n)					
<15	6.1	2.6	3.0	4.7	1.3	17.7
15–19	1.8	1.6	2.3	2.6	1.8	10.1
≥20	1.1	1.4	2.8	3.1	2.3	10.7
Total	9.0	5.6	8.2	10.3	5.4	38.5
	F, *Total*					
<15	42.2	17.5	20.8	32.4	9.4	122.3
15–19	13.6	11.9	17.3	19.3	12.7	74.8
≥20	11.2	13.1	25.4	27.7	20.0	97.3
Total	67.0	42.5	63.5	79.4	42.1	294.5
	G, *As a percentage of the total sample*					
<15	2.9%	1.2%	1.4%	2.2%	0.6%	8.4%
15–19	0.9	0.8	1.2	1.3	0.9	5.2
≥20	0.8	0.9	1.8	1.9	1.4	6.7
Total	4.6%	2.9%	4.4%	5.5%	2.9%	20.3%

with no experience could have left question 9 blank as a self-evident "no," possibly failing to notice or ignoring the instruction to write "never" in the blank space labeled "age." In addition, some respondents may have found the instruction confusing.

Measures of Reliability

Several summary measures of interest from the data in Table 2, including estimates for the marginal distribution of frequency of contact; estimates for the percentage of the sample with contact at age 15 or older, and at age 20 or older are shown in Table 3; and, within these groups, respondents are further classified by the reported frequency of homosexual contact. Table 3 also provides measures of reliability for these estimates, specifically estimates of standard errors based on the jackknife method (*22*). These standard errors reflect the effect of both the sample design and the random variation arising from estimating distributions for incomplete cases from the complete cases. The standard errors were estimated on the basis of variation among 73 sets of estimates, each derived by omitting one of the 73 primary sampling units and carrying through the missing data analysis on the reduced sample. Because the actual sample included self-representing primary sampling units and stratification, the calculation actually overstates the variability due to sampling error, although this effect appears to be slight.

For purposes of comparison, Table 3 presents the standard errors that would apply if these estimates were derived from a simple random sample of the same size with no missing data. The estimates are less reliable, but not dramatically less reliable, than would be the case if they had been derived from a simple random sample (with no missing data).

Estimates derived exclusively from cases with complete data are also shown in Table 3 (*23*). We observe that failure to impute responses for cases with incomplete data yields estimates for age at last contact and frequency of contact that are generally approximately a third lower than estimates derived from the missing data analysis, primarily because the imputation includes an allowance for several incomplete cases indicating some homosexual contact.

Table 3. Comparisons of estimates that include imputation for missing data with estimates based only on cases with complete responses to all questions.

Frequency of activity	Total sample including imputations (%)			Complete cases only (%)	
Any homosexual experience	20.3%	(1.2)*	(1.1)†	14.4%	(1.3)*
1 Once	4.6	(0.6)	(0.6)	3.5	(0.5)
2 Twice	2.9	(0.7)	(0.4)	2.1	(0.6)
3 Rarely	4.4	(0.6)	(0.5)	3.1	(0.5)
4 Occasionally	5.5	(0.9)	(0.6)	3.9	(0.7)
5 Fairly often	2.9	(0.6)	(0.4)	1.9	(0.4)
Levels 3–5	12.8	(1.3)	(0.9)	8.8	(1.0)
Levels 4–5	8.4	(1.1)	(0.7)	5.7	(0.8)
Last contact at ≥15 years	11.9	(1.0)	(0.9)	7.8	(0.9)
Levels 3–5	8.4	(1.0)	(0.7)	5.5	(0.7)
Levels 4–5	5.5	(0.9)	(0.6)	3.5	(0.6)
Level 5	2.3	(0.5)	(0.4)	1.4	(0.4)
Last contact at ≥20 years	6.7	(0.8)	(0.7)	4.0	(0.6)
Levels 3–5	5.0	(0.8)	(0.6)	3.0	(0.6)
Levels 4–5	3.3	(0.7)	(0.5)	1.9	(0.4)
Level 5	1.4	(0.4)	(0.3)	0.8	(0.2)

*Standard errors based on complex design and, in the case of the estimates that include imputations, calculations include effects of random variation in estimating the missing data. †Standard errors computed under the assumption of simple random sampling for a sample of the same size, that is, 1450 cases.

Estimates for Subgroups

The relative similarity of the distributions of same-gender sexual experiences across age groups is striking (Table 4). The major difference observed appears to be the lower reporting of experience by men 65 years or older. This difference appears largely to be due to the lower proportion of older men reporting activity that ended by age 15.

The breakdown by education (Table 5) contains several interesting features. Men with four or more years of college are estimated to have a higher proportion with experience, particularly compared to those with no college education. Much of this difference appears to arise from adolescent and preadolescent activity, however. The estimates for the highest education group are also generally the least

Table 4. Estimates of homosexual experience by age of respondent at time of interview.

Frequency of activity	Age 21–29		Age 30–44		Age 45–64		Age 65+	
	Reported and imputed (%) (n = 270)	Only reported (%) (n = 221)	Reported and imputed (%) (n = 452)	Only reported (%) (n = 362)	Reported and imputed (%) (n = 457)	Only reported (%) (n = 338)	Reported and imputed (%) (n = 271)	Only reported (%) (n = 214)
Any experience	21.6 (2.6)*	15.4	22.3 (2.0)*	17.7	21.5 (2.0)*	14.8	13.6 (2.0)*	7.5
1 Once	6.0 (1.3)	5.0	5.5 (1.0)	5.0	3.8 (0.8)	2.4	3.0 (0.8)	1.4
2 Twice	3.8 (1.2)	3.2	2.8 (0.9)	2.2	3.0 (0.8)	1.8	2.2 (0.8)	1.4
3 Rarely	4.1 (1.1)	2.7	5.5 (1.0)	4.4	4.7 (1.0)	3.3	2.4 (0.6)	0.9
4 Occasionally	4.9 (1.1)	3.2	5.6 (1.3)	4.1	6.4 (1.2)	4.7	4.2 (1.2)	2.8
5 Fairly Often	2.7 (1.1)	1.4	3.0 (0.8)	1.9	3.6 (0.9)	2.7	1.9 (0.6)	0.9
Levels 3–5	11.7 (2.2)	7.2	14.0 (1.7)	10.5	14.6 (1.8)	10.7	8.5 (1.6)	4.7
Levels 4–5	7.6 (1.7)	4.5	8.6 (1.5)	6.1	10.0 (1.5)	7.4	6.1 (1.6)	3.7
Last contact ≥15	11.2 (2.1)	7.2	12.2 (1.6)	9.1	13.6 (1.5)	8.3	9.2 (1.7)	5.1
Levels 3–5	7.5 (1.9)	4.5	8.6 (1.2)	6.1	10.0 (1.4)	6.5	6.6 (1.4)	3.7
Levels 4–5	4.8 (1.4)	2.7	5.1 (0.9)	3.3	6.9 (1.2)	4.7	4.6 (1.4)	2.8
Level 5	2.3 (1.0)	1.4	2.2 (0.7)	1.4	2.7 (0.8)	1.8	1.7 (0.6)	0.9
Last contact ≥20	6.5 (1.5)	4.1	6.0 (1.1)	3.9	8.1 (1.2)	4.7	5.7 (1.4)	2.8
Levels 3–5	5.0 (1.5)	3.2	4.3 (0.9)	2.5	6.4 (1.2)	4.1	4.1 (1.0)	1.9
Levels 4–5	3.1 (1.1)	1.8	3.0 (0.7)	1.9	3.9 (0.9)	2.4	2.8 (1.0)	1.4
Level 5	1.5 (0.6)	0.9	1.1 (0.4)	0.6	1.8 (0.6)	1.2	1.1 (0.5)	0.5

*Numbers in parentheses are standard errors that were calculated to take account of the complex sample design and the effects of random variation in imputing the missing data.

dependent on the imputation for missing data because this group had the most complete reporting. For example, completely reported cases account for 77 percent of the estimated 31.8 percent of college graduates with some homosexual experience (*24*), whereas they account for only 44 percent of the estimated 17.4 percent of those with less than a high school education.

Table 6 shows the marital status of respondents classified as currently married; divorced, widowed, or separated; or single (that is, never married). Single men are further divided by age into 21 to 29 versus ≥30 years old. Although the sample sizes are small and the reliability less than would be desirable, single men age 30 and over appear the most likely of the four groups to have had same-gender sexual experiences. For this group, the differences appear to arise not only from adolescent activity but also from contacts after age 20 as well.

Comparison of Estimates from 1970 and 1988 Surveys

Michael *et al.* (7) have reported a rate of 3.2 percent for homosexual contact during the past 12 months among men (≥18) reporting any partners of either sex (including wives) during this period. When men reporting no partners are included in the

Table 5. Estimates of homosexual experience by education.

Frequency of activity	Less than high school		High school		Some college		College graduate or higher	
	Reported and imputed (%) (n = 619)	Only reported (%) (n = 473)	Reported and imputed (%) (n = 395)	Only reported (%) (n = 303)	Reported and imputed (%) (n = 237)	Only reported (%) (n = 189)	Reported and imputed (%) (n = 195)	Only reported (%) (n = 167)
Any experience	17.4 (1.6)*	10.1	19.6 (1.8)*	12.2	20.0 (2.9)*	16.4	31.8 (3.7)*	28.7
1 Once	3.6 (0.7)	1.9	4.4 (0.9)	3.0	4.8 (1.1)	4.2	8.1 (1.7)	8.4
2 Twice	3.3 (0.8)	2.5	1.7 (0.5)	0.3	2.6 (0.9)	2.1	4.7 (1.7)	4.2
3 Rarely	3.2 (0.6)	1.5	5.5 (1.1)	4.3	4.8 (1.1)	4.2	5.5 (1.4)	4.2
4 Occasionally	4.4 (0.8)	2.5	5.5 (1.2)	3.6	4.7 (1.6)	3.7	9.8 (2.4)	8.4
5 Fairly often	2.9 (0.7)	1.7	2.5 (0.7)	1.0	3.0 (1.2)	2.1	3.6 (1.2)	3.6
Levels 3–5	10.5 (1.2)	5.7	13.4 (1.8)	8.9	12.5 (2.6)	10.1	18.9 (3.0)	16.2
Levels 4–5	7.3 (1.1)	4.2	8.0 (1.5)	4.6	7.7 (2.3)	5.8	13.5 (2.7)	12.0
Last contact ≥15	11.7 (1.3)	6.8	11.1 (1.3)	6.3	10.4 (2.5)	7.9	16.1 (2.5)	13.2
Levels 3–5	8.0 (1.0)	4.4	8.3 (1.3)	5.0	6.7 (1.9)	4.8	12.3 (2.3)	10.2
Levels 4–5	5.8 (1.0)	3.6	4.5 (1.0)	2.0	5.0 (1.5)	3.7	7.1 (2.0)	6.0
Level 5	2.5 (0.7)	1.5	1.6 (0.5)	0.3	2.7 (1.2)	2.1	2.5 (1.1)	2.4
Last contact ≥20	6.4 (1.0)	3.0	6.8 (1.2)	3.6	6.2 (1.8)	4.8	8.3 (1.8)	6.6
Levels 3–5	4.8 (0.9)	2.3	5.3 (1.1)	3.0	4.3 (1.7)	3.2	6.1 (1.6)	4.8
Levels 4–5	3.5 (0.8)	1.9	2.5 (0.7)	0.7	3.4 (1.3)	2.6	4.2 (1.4)	3.6
Level 5	1.4 (0.5)	0.6	1.1 (0.4)	0.3	1.4 (0.7)	1.1	1.8 (0.9)	1.8

*Standard errors were calculated to take account of the complex sample design and the effects of random variation in imputing the missing data.

Table 6. Estimates of the percentage of the adult male population with homosexual experience by marital status, age at last contact, and frequency of contact during the period of peak activity.

Frequency of activity	Currently married		Divorced, widowed or separated		Single, age 21–29		Single, age ≥30	
	Reported and imputed (%) (n = 1161)	Only reported (%) (n = 899)	Reported and imputed (%) (n = 154)	Only reported (%) (n = 127)	Reported and imputed (%) (n = 80)	Only reported (%) (n = 68)	Reported and imputed (%) (n = 55)	Only reported (%) (n = 41)
Any experience	19.6 (1.4)*	14.1	19.3 (3.2)*	13.4	23.7 (4.1)*	16.2	32.9 (7.7)*	22.0
1 Once	4.7 (0.6)	3.8	3.4 (1.3)	1.6	4.4 (1.9)	2.9	6.4 (2.2)	4.9
2 Twice	2.5 (0.7)	1.7	3.5 (1.4)	2.4	5.0 (2.9)	4.4	7.3 (4.3)	7.3
3 Rarely	4.4 (0.7)	3.1	3.3 (1.0)	2.4	5.7 (2.2)	4.4	4.7 (2.2)	2.4
4 Occasionally	5.4 (0.9)	4.0	5.0 (1.6)	3.1	4.9 (2.1)	2.9	9.0 (4.1)	4.9
5 Fairly often	2.6 (0.5)	1.6	4.1 (1.6)	3.9	3.8 (1.9)	1.5	5.5 (2.8)	2.4
Levels 3–5	12.4 (1.3)	8.7	12.4 (2.8)	9.4	14.3 (3.6)	8.8	19.2 (5.8)	9.8
Levels 4–5	8.0 (1.1)	5.6	9.1 (2.5)	7.1	8.7 (2.9)	4.4	14.5 (5.2)	7.3
Last contact ≥15	10.9 (0.9)	6.8	13.2 (2.8)	10.2	15.1 (3.8)	10.3	23.1 (5.7)	17.1
Levels 3–5	7.7 (0.9)	4.7	10.6 (2.6)	9.4	9.7 (3.0)	5.9	15.6 (5.1)	9.8
Levels 4–5	4.9 (0.8)	2.9	7.8 (2.4)	7.1	5.9 (2.3)	2.9	12.0 (4.7)	7.3
Level 5	1.9 (0.5)	1.0	3.9 (1.7)	3.9	3.2 (1.6)	1.5	4.5 (2.5)	2.4
Last contact ≥20	6.0 (0.8)	3.1	7.2 (2.0)	5.5	10.1 (2.9)	7.4	15.2 (4.7)	12.2
Levels 3–5	4.5 (0.7)	2.3	5.6 (1.8)	4.7	6.6 (2.5)	4.4	12.3 (4.7)	9.8
Levels 4–5	2.8 (0.7)	1.4	4.4 (1.4)	3.9	3.3 (1.7)	1.5	9.3 (4.2)	7.3
Level 5	1.2 (0.4)	0.6	1.7 (0.9)	1.6	2.4 (1.5)	1.5	3.5 (2.2)	2.4

*Standard errors were calculated to take account of the complex sample design and the effects of random variation in imputing the missing data.

denominator, the rate for all men 21 years or older becomes 3.0 percent. Since these data were derived from an NORC household sample (*25*), one further adjustment is appropriate. In the NORC General Social Survey, one respondent age 18 or more is randomly selected within each sampled household. Thus adults living with other adults have a lower probability of selection than adults who live alone. Weighting the observations by the number of persons age 18+ in the household would compensate for the differential probabilities of selection within households.

For purposes of comparison, Table 7 provides estimates of the proportion of this 1988 sample reporting same-gender sexual contact in the last 12 months. The overall weighted estimate (2.4 percent) is smaller than its unweighted counterpart (3.0 percent) because married men, who typically live in households with two or more eligible adults, receive a higher than average weight and have a lower than average probability of reporting homosexual contact within the last 12 months.

Also for purposes of comparison, Table 7 displays two measures from the 1970 Kinsey study approximating the concept measured in the General Social Survey: (i) the proportion of men whose age at last homosexual contact was the same as their current age and (ii) the proportion with age at last homosexual contact within 1 year of their current age. These two measures represent lower and upper bounds on contact during the previous 12 months. The first and fourth columns report estimates adjusted for missing data. The estimates are derived by combining, separately by marital status, (i) the information provided by the subset of respondents fully answering the questions on frequency of contact and age at last contact, and (ii) the overall estimates by frequency of contact, including imputations, for the number of men with an age of last contact of 20 or older (*26*). The adjusted estimates from this process are shown both by marital status and by age in Table 7.

To show the effects of the missing data adjustments, Table 7 also displays the respondents with reported recent contact as percentages of all cases, including nonrespondents. The comparison shows that a substantial proportion of the estimates arises from our adjustments for missing data, especially for currently married men.

With relatively large sampling errors affecting the estimates from both sources, the 1970 Kinsey and 1988 General Social Survey estimates appear close to each other. Both sources also point to a particularly significant contribution of currently married men to homosexual activity although there are reasons to question the latter finding with respect to the 1988 General Social Survey (*27*).

Further corroborating the results indicated in Table 7, an analogous analysis of the missing data on age at last contact yields an estimate that approximately 1.1 percent (SE, 0.4) of all currently married men had a last homosexual contact at the same age as their first marriage, and an additional 3.7 percent (SE, 0.6) had homosexual contact at ages beyond the age of first marriage. (The estimated rates for those with only one marriage are 1.3 percent and 3.0 percent, respectively, showing that the overall result for currently married men is not simply due to men with more than one marriage.) We note again that approximately half or more of these estimated percentages arise from our imputation of the missing data.

Number of Partners

In Table 8 are estimates of the number of lifetime male partners, reflecting weighting adjustments for missing data similar to those used to construct the estimates (Table 7) from the 1970 Kinsey study. The observed distribution of reported partners was used to adjust the overall estimates of same-gender sexual experience by marital status, age at last contact, and frequency of contact. These adjustments suggest that the number of men with five or more male partners is a relatively small fraction of the total number of men who have had some sexual experience with other males. An estimated 2.1 percent (SE, 0.7) of the male population had both sexual contact with a man after age 20 and also had five or more male partners during his lifetime; 1.2 percent (SE, 0.5) had 10 or more male partners. We note that of the 15 fully observed cases meeting the last condition, that is, with contact after age 20 and with 10 or more lifetime partners, there are six responses of 10, four of 20, one each of 15, 19, and 50, and two respondents coded as having 100 or more male partners.

We also compared responses on numbers of homosexual partners, without adjustments for missing data, to the observed distributions for numbers of female partners before marriage (*28*) collected in an earlier series of questions included on the self-administered questionnaire. Although these data are sparse, they suggest that currently or previously married men with same-gender sexual experience may, on average, have had approximately as many female partners as exclusively heterosexual men. For example, there are 11 currently or

Table 7. Proportion (and standard errors) of men with reported or imputed recent homosexual contact from the 1970 Kinsey Survey and from the 1988 NORC General Social Survey, (GSS).

Age and marital status	1970 Kinsey Survey						1988 NORC-GSS	
	Contact since last birthday*			Since preceding birthday†			During last 12 mos.§	
	Reported and imputed	SE	Reported‡ only	Reported and imputed	SE	Reported‡ only	Reported§ only	SE
All men	1.6	(0.5)	0.8	2.0	(0.5)	1.0	2.0	(0.7)
Marital status								
Currently married	1.3	(0.5)	0.5	1.5	(0.6)	0.6	1.5	(0.8)
Div., Wid., Sep.	0.9	(0.9)	0.6	1.8	(1.2)	1.3	2.2	(1.8)
Never married	5.3	(2.4)	3.7	6.2	(2.6)	4.4	3.0	(1.7)
Age								
21–29	3.4	(1.6)	1.9	4.8	(1.8)	2.6	3.8	(1.1)
30–44	0.9	(0.6)	0.4	1.1	(0.6)	0.7	2.1	(0.9)
45–64	2.3	(1.1)	1.1	2.3	(1.1)	1.1	2.2	(1.3)
≥65	0.0		0.0	0.0		0.0	2.9	(1.4)

*Respondents with age of last homosexual contact the same as the current age. †Respondents with age of last homosexual contact the same or one year less than the current age. ‡The base of each percentage is the total number of sample persons, including nonrespondents. §The base of each percentage is the number of respondents, excluding nonrespondents. This sample was restricted to men 21 years of age and older; tabulations were weighted to adjust for different probabilities of selection for persons residing in households with more than one eligible adult.

previously married men with reported homosexual contact at age 20 or older and five or more male partners. Ten of these 11 men report five or more female partners. Similarly the reporting of heterosexual contact among never married men is substantial. Significantly, five of 12 never married men reporting homosexual contact at age 20 years or older reported having five or more female partners (with three others not responding).

Sensitivity of Estimates to Assumptions

Although the standard errors we report in Tables 2 through 8 measure, or slightly overstate, the effect of random variation on our inferences, they do not represent the uncertainty introduced by the choice of assumptions used in the imputation of missing data. To suggest the degree to which our results are sensitive to this choice, we present two additional sets of estimates that are based on different treatments of the missing data.

The principal set of estimates (Tables 2 through 6), and the related analyses (Tables 7 and 8), employ observed data to the extent possible in estimating the missing data. In fact, the estimates represent a practical approximation to the model of ignorable nonresponse, for which it is assumed (conditional on the values of the observed data) that the probabilities of nonresponse do not depend on the unobserved data. In other words, nonrespondents were assumed similar to respondents with the same reported information. Because of the patterns of response in this application, maximum likelihood estimation of the ignorable model would require iterative calculations. Practically speaking, however, the results would substantially agree with those presented in Tables 2 through 8.

As one alternative, we present the outcome of a simpler procedure for handling the missing data; it is based on use of only the distributions of the 1135 cases with complete data on both variables—that is, those with reports of no homosexual experience and those whose frequency of contact and age at last contact are both reported. For this alternative we imputed values for missing data as follows: (i) those missing age at last contact only were imputed to the overall distribution of this variable for those with homosexual contact among the complete cases; (ii) those missing frequency of contact only were imputed to the overall distribution of this variable for those with homosexual contact among the complete cases; (iii) those missing both variables but with reported homosexual activity were imputed to the distribution for complete cases with homosexual contact; and (iv) those with completely missing data were imputed to the overall distribution of complete cases. This alternative imputation strategy mimics the more standard practices of survey research when insufficient resources are available for a focused analysis of this one missing data problem.

A second alternative incorporates explicit assumptions that the probability of response depended on values of the missing data. The specific assumptions for this alternative derive from two conjectures. First, we conjecture that respondents may have been more inclined to answer a question on frequency of contact if the actual answer could be given as "only once," or "twice," rather than requiring the respondent to make a classification into the vaguer categories of "rarely," "occasionally," or "fairly often." Second, we conjecture that respondents currently homosexually active may have misinterpreted the sense of "your age the last time you had this experience," if "last" suggested to them a cessation. Hence, those currently active may have been less inclined to answer this question than those whose activities ended in their teens.

To illustrate the implications of these two conjectures, a second alternative set of estimates was derived according to the following procedures: (i) those missing age at last contact only were imputed to ages of 20 years and older; (ii) those missing frequency of contact only were imputed to the conditional distribution over the upper three categories of frequency of homosexual contact among the complete cases, according to age at last contact; (iii) those missing both variables but with reported homosexual activity were imputed to age at last contact as 20 or older and to the conditional distribution of frequency of contact among the upper three categories, as in the preceding step; and (iv) those with completely missing data were imputed to the distribution of the remaining cases, including cases with distributions imputed in the preceding three steps. This procedure reflects an assumption that nonresponse is nonignorable, that is, the probabilities of response depend directly on the values of the missing data. The procedure closely resembles one specific missing data model (*29*) from a class of nonignorable models for categorical data.

In Table 9, results from our principal analysis of the missing data are compared with the two alternatives. Differences between the principal analysis and the first alternative are slight, although the first alternative shows a slightly lower level of frequency of contact

Table 8. Estimates of the distribution of men with homosexual contact by the number of male partners and age at last homosexual contact. Men reporting only one homosexual contact are included as having one partner. Standard errors are shown in parentheses.

Group	Number of male partners (%)				
	1	2–4	5–9	≥10	≥5
	As a percentage of overall male population				
All men, last contact	7.3 (0.8)	9.3 (1.0)	1.6 (0.4)	2.1 (0.7)	3.7 (0.9)
<15 years*	4.1 (0.7)	3.7 (0.7)	0.3 (0.2)	0.1 (0.1)	0.4 (0.2)
15–19 years*	1.6 (0.4)	2.5 (0.6)	0.4 (0.2)	0.8 (0.4)	1.2 (0.5)
≥20 years*	1.6 (0.4)	3.0 (0.6)	0.9 (0.3)	1.2 (0.5)	2.1 (0.6)
	With last contact ≥20 (% of male population)				
Marital status					
Currently married	1.5 (0.5)	3.1 (0.7)	0.4 (0.3)	1.1 (0.5)	1.5 (0.5)
Divorced, widowed, or separated	2.0 (1.4)	0.0	4.2 (1.6)	1.1 (1.1)	5.3 (2.0)
Never married	2.7 (1.7)	5.6 (2.1)	2.2 (1.6)	1.8 (1.3)	3.9 (1.9)
	With last contact ≥20 (% of male population)				
Age (years)					
21–29	1.6 (1.2)	4.2 (2.1)	0.5 (0.5)	0.9 (0.9)	1.4 (1.0)
30–44	1.0 (0.4)	3.1 (1.1)	1.1 (0.6)	1.2 (0.9)	2.3 (1.1)
45–64	2.8 (1.0)	2.8 (1.2)	1.3 (0.8)	1.9 (0.9)	3.2 (1.3)
≥65	0.8 (0.5)	1.8 (1.3)	0.6 (0.6)	0.0	0.6 (0.6)

*The base of the percentages is the total male population, so that percentages on the three lines sum to the percentages on the first line, except for differences due to rounding.

Table 9. Comparison of estimates of the percentage of the adult male population with homosexual experience under two alternative assumptions.

Frequency of activity	Estimates (%)		
	Primary analysis	Alternative 1	Alternative 2
Any homosexual experience	20.3	19.9	21.1
1 Once	4.6	4.7	3.5
2 Twice	2.9	2.9	2.1
3 Rarely	4.4	4.2	5.4
4 Occasionally	5.5	5.3	6.5
5 Fairly often	2.9	2.7	3.7
Levels 3–5	12.8	12.3	15.5
Levels 4–5	8.4	8.0	10.1
Last contact at age ≥15	11.9	11.2	14.4
Levels 3–5	8.4	7.8	11.9
Levels 4–5	5.5	5.0	7.8
Level 5	2.3	2.0	3.2
Last contact at age ≥20	6.7	6.4	10.6
Levels 3–5	5.0	4.6	9.4
Levels 4–5	3.3	3.0	6.2
Level 5	1.4	1.2	2.6

and proportion active at older ages. The second alternative shows much higher estimates for high frequency of contact and for contact after age 20; in fact, the estimates more than double for those with activity after age 20 and more than two homosexual contacts. The second alternative would imply increases in the estimated numbers of men with recent contact in Table 7 and numbers of male partners presented earlier. The estimated total number of men with homosexual experience is relatively close to that found in the primary analysis, however.

The second alternative may in one sense be characterized as more extreme than the first alternative or the primary imputation, because the second alternative treats as impossible some events that the other two analyses allow. For example, only the second alternative assumes that every respondent whose actual experience ended before age 20 will answer the question on age of last contact whenever the respondent answers the question on frequency of contact. In practice, it would be highly unusual for nonresponse to follow such a rigid pattern. Nonetheless, the second alternative suggests the possibility that our principal estimates may understate activity at older ages and higher frequencies due to the sensitivity of our estimates to the assumptions made in our imputation of missing data.

Conclusions

There are unavoidable problems in developing and interpreting estimates derived from survey data such as these. Most importantly, the available data do not allow any independent validation of the accuracy of the men's self-reports. Given a history of discrimination and oppression—not to mention the fact that in many states sodomy statutes would classify the reported behaviors as criminal—it is reasonable to expect that there will be a downward bias in estimates derived from self-reports. That is to say, it is reasonable to assume that more men will falsely report the absence of same-gender sexual experiences than will report experiences that never occurred.

In our analysis of complete cases, reports of same-gender sexual experience appeared to be strongly associated with responses to a number of attitude items and being unmarried beyond age 30. We have used this evidence and the responses to the attitude items by those not responding to the homosexual experience question to conclude that the 223 complete nonrespondents, as a group, did not have a higher proportion with same-gender sexual experience than the rest of the sample. We recognize the possibility that some respondents, attempting to deny or conceal a past history, may have both skipped the self-administered questions and held or reported attitudes considerably less liberal about sexual matters than other respondents with homosexual experience who were willing to report this experience on a self-administered questionnaire. It is also possible that respondents with deeply held views against homosexuality and other strongly conservative opinions on sexual behavior might willingly report these on the interviewer-administered section of the interview, yet disproportionately deny homosexual activities of their own on the self-administered questionnaire by skipping these items.

With regard to the sample design, our confidence in the adequacy of the 1970 Kinsey-NORC probability sample with quotas is enhanced by the rough agreement found between estimates derived from that sample and estimates obtained from the 1988 NORC–General Social Survey sample, which employed a full-probability design. We would point out, however, that sample designs of both these surveys exclude the population of men residing in institutions or group quarters (including men in prison or military barracks). Same-gender sexual contact may be more common in some of these settings (*30*).

Our analyses do point to a relative complexity and variety in the population with homosexual experience with respect to timing and frequency of contact. Clearly, a substantial number of respondents were willing to report on these activities on a self-administered questionnaire under the conditions of the survey. Thus, our analyses indicate that roughly one-fifth of adult American males (in 1970) had at least one homosexual experience (where one partner reached orgasm). This proportion is lower than that published by Kinsey, who found, "In these terms of (physical contact to the point of orgasm), the data in the present study indicate that at least 37 percent of the male population had some homosexual experience . . ." (*2*, p. 623).

Because of the oddities of the original Kinsey sample, the new figure may provide a better reference point. However, given the response biases that one can reasonably assume to operate, this new figure might be taken as a lower bound. Clearly, there is good reason to suspect that the net bias in self-reports of homosexual experiences is negative (as noted above). Societal intolerance of same-gender sex may diminish survey respondents' willingness to provide complete and accurate reports of behaviors that are classified as crimes in many states. The uncertainties introduced by this factor will bedevil future attempts to estimate the prevalence of same-gender sexual behaviors and to infer changes across times in the prevalence and patterns of these behaviors. Indeed, it is reasonable to expect that the magnitude of the reporting bias will vary with the levels of societal intolerance. Willingness to report homosexual behaviors in a survey is likely to vary over time in response to many of the same social forces that are likely to influence the behavior itself, thereby creating a most difficult (and possibly intractable) measurement problem.

In considering the epidemiology of AIDS, it is reasonable to focus on adult experiences that occurred with some frequency rather than on rare experiences confined entirely to adolescence. In that regard, it is estimated that 1.4 percent of men had adult homosexual contacts (for example, at age 20 years and older) whose frequency was characterized as being "fairly often" (at some point in time). An additional 1.9 percent of men had adult experiences whose frequency was characterized as "occasionally." Taken together, these two groups made up 3.3 percent of the adult male population. Here again, the nature of the presumed reporting bias prompts the caveat that this is likely to be a lower bound estimate of the true proportion

of men in the population who have this level of same-gender sexual experience. Furthermore, our exploration of the sensitivity of our conclusions to variations in assumptions (Table 9) indicates that the number might reasonably be estimated at almost twice this percentage (6.2 percent, Table 9).

Overall, these numbers appear similar to the 1948 Kinsey estimate used by the Public Health Service in its projections (that is, that 4 percent of U.S. men are "exclusively homosexual" throughout their lives). In fact, the interpretation of our estimates is different. Most of the men included in our 3.3 percent estimate could not be classified as "exclusively homosexual" throughout their lives. Thus, as Table 6 indicates, men who are currently or were previously married (who are a much larger segment of the total population) account for the majority of men who would be included in our estimate that 3.3 percent of adult men in 1970 had same-gender sexual contacts during adulthood whose frequency (at their peak) was characterized as "fairly often" (1.4 percent) or "occasional" (1.9 percent). Beyond the findings reported here from these two surveys, the problems of analysis and interpretation that we have encountered suggest the need for continuing research to improve the design and application of surveys on sexual behavior in order to reduce problems of misinterpretation, misreporting, and nonresponse.

REFERENCES AND NOTES

1. U.S. Public Health Service, *Public Health Rep.* **101**, 341 (1986).
2. A. C. Kinsey, W. B. Pomeroy, and C. E. Martin [*Sexual Behavior in the Human Male* (Saunders, Philadelphia, 1949)] conclude (p. 650) that "37 percent of the total male population has at least some overt homosexual experience to the point of orgasm between adolescence and old age." They also conclude (p. 651) that "10 percent of the males are more or less exclusively homosexual (that is, rate 5 or 6) for at least 3 years between the ages of 16 and 55 . . . 8 percent of the males are exclusively homosexual (that is, rate 6) for at least three years between the ages of 16 and 55"; and that "4 percent of the white males are exclusively homosexual throughout their lives after the onset of adolescence." Kinsey used a heterosexual-homosexual scale with ratings from 0 to 6 that was based on both psychological reactions and overt experience," with 0 being "exclusively heterosexual, no homosexual" and 6 "exclusively homosexual, no heterosexual" (p. 638).
3. Centers for Disease Control, "Human immunodeficiency virus infection in the United States: A review of current knowledge and plans for expansion of HIV surveillance activities," Report to the Domestic Policy Council (Department of Health and Human Services, Washington, DC, November 1987), table 14.
4. See, for example, W. A. Wallis, *J.Am. Stat. Assoc.* **44**, 463 (1948); L. M. Terman, *Psychol. Bull.* **45**, 443 (1948); W. G. Cochran, F. Mosteller, J. W. Tukey, *J.Am. Stat. Assoc.* **48**, 673 (1948).
5. The timing of Kinsey *et al.*'s work virtually coincided with that of J. Neymann on representative methods in population sampling [*J. R. Stat. Soc.* **A97**, 558 (1934); *J. Am. Stat. Assoc.* **33**, 101 (1938)].
6. Some further data are available although they do not satisfy our need for data on sexual behavior in representative samples of the population. A re-tabulation of the Kinsey data (including data collected through 1963) is reported in P. H. Gebhard and A. B. Johnson [*The Kinsey Data: Marginal Tabulations of the 1938–1963 Interviews Conducted by the Institute for Sex Research* (Saunders, Philadelphia, 1979)]. Data are presented from a "basic sample" of "postpubertal individuals who were never convicted of any offense other than traffic violations and who did not come from any source which we knew to be biased in terms of sexual behavior (p. 41)," thereby eliminating persons recruited from prisons and groups such as the Mattachine Society. For this "basic sample," Gebhard and Johnson (p. 428, table 379) report that 36.3 percent of white college-educated males, 39.8 percent of white noncollege-educated males, and 30.0 percent of black males had some homosexual experience (including 0.6 to 2.1 percent reporting arousal without physical contact). In this tabulation, Gebhard and Johnson find that between 9.0 and 12.7 percent of the male sample report "extensive" experience, which is defined as ≥51 homosexual experiences, ≥21 partners, or both. Reviewing studies by Kinsey and others published through 1967, Gebhard concluded that "all in all, at our present state of knowledge, I think we can say that between one-quarter and one-third of the adult males of the college-educated segment of the population have had overt homosexual experience since puberty and that most of this was confined to adolescence" [P. H. Gebhard, in *NIMH Task Force on Homosexuality, Final Report and Background Papers* (U.S. Government Printing Office, Washington, DC, 1972), p. 27]. Gebhard further concludes that "about 4 percent of white college-educated adult males are predominantly homosexual" and that "among the unmarried, exclusive homosexuality (6 on the scale) is commoner than predominant homosexuality (4 or 5) so that the bulk of this 4 percent are exclusively homosexual" (*ibid.*, p. 28) [see (*2*) for a description of the Kinsey rating scale]. After evaluating the available data, the NIMH task force charged to "review carefully the current state of knowledge regarding homosexuality" recommended that collection of better data on the "incidence and epidemiology of homosexual experience and behavior" should be given high priority because "present estimates are based on data which are inadequate and out of date" (p. 3). Some studies published since 1969 have reported estimates of the prevalence of homosexual contact in large samples [see, for example, M. Hunt, *Sexual Behavior in the 1970s* (Playboy Press, Chicago, 1974)]. However, the research designs of the studies of which we are aware do not permit confident generalization to the national population.
7. R. T. Michael *et al.*, *Morb. Mortal. Weekly Rep.* **37**, 565 (1988).
8. Surveys of convenience samples of members of homosexual or gay communities report high rates of robbery, blackmail, job loss, discharge from the military, and social ostracism [W. Simon and J. Gagnon, *J. Health Soc. Behav.* **8**, 177 (1967); A. P. Bell and M. S. Weinberg, in *Homosexualities: A Study of Diversity Among Men and Women* (Simon & Schuster, New York, 1978)]. For a historical perspective on the harrassment of gay men and lesbians, see J. D'Emilio in *Sexual Politics, Sexual Communities: The Making of a Homosexual Minority in the United States* (Univ. of Chicago Press, Chicago, 1983); J. N. Katz, *Gay/Lesbian Almanac: A New Documentary* (Harper and Row, New York, 1983). See also C. F. Turner *et al.*, *AIDS, Sexual Behavior, and IV Drug Use* (National Academy Press, Washington, DC, in press), figure 7-2.
9. For further discussion of the effects of "social desirability" and survey response, see T. J. DeMaio, in *Surveying Subjective Phenomena*, C. F. Turner and E. Martin, Eds. (Russell Sage-Basic Books, New York, 1984); N. M. Bradburn and S. Sudman, *Improving Interview Method and Questionnaire Decision* (Jossey-Bass, San Francisco, 1979).
10. E. E. Levitt and A. D. Klassen, *J. Homosexuality* **1**, 29 (1974); A. D. Klassen, C. J. Williams, E. E. Levitt, *Sex and Morality in the United States*, H. J. O'Gorman, Ed. (Wesleyan Univ. Press, Middleton, CT, in press).
11. W. Booth, *Science* **239**, 1084 (1988); A. D. Klassen, *ibid.*, **240**, 375 (1988).
12. It should be recognized, however, that the use of quota sampling at the block level provides a less secure basis for inference than a full probability-based design.
13. This rate is based on nonresponse to one or both of the questions analyzed in detail in Tables 2 through 6. There was additional nonresponse to other items in this section of the questionnaire, including items analyzed in Tables 7 and 8.
14. National Opinion Research Center, "The sample design for NORC's amalgam surveys," unpublished memorandum from NORC Sampling Department, University of Chicago, February 1976.
15. In this design, which has been called "probability sampling with quotas (PSQ)," controls are imposed on interviewer discretion in choosing the household to be interviewed. This design was used in the survey reported here and also in the 1972, 1973, and 1974 waves of NORC's General Social Survey (the National Data Program for the Social Sciences). In the 1975 and 1976 surveys, an experiment was carried out as part of that survey's change-over to a full probability sample design with random selection at the block level [C. B. Stephenson, *Public Opinion Q.* **43**, 477 (1979)]. In this experiment, half of each survey sample in 1975 and 1976 was followed with full probability methods and half with the earlier design of "probability sampling with quotas." In this experiment, the probability sampling with quotas (PSQ) sample established quotas similar to that in the 1970 surveys, the quotas being for men under 35, men 35 and older, employed women, and women who were not employed. Generally, the agreement between survey results was quite close except that the PSQ sample included somewhat more large households than the full probability method. Both PSQ and full probability methods, however, underrepresent persons from large households (when data are analyzed without use of household sampling weights). The result of these two processes yielded the following estimates of the proportion of the population living in households of different sizes (full probability first; then PSQ estimates): one in 1-adult households, .12 and .15; in 2-adult households, .57 and .60; in 3-adult households, .18 and .17; and in ≥4-adult households, .11 and .08. (These figures are averages of estimates reported from 1975 and 1976 surveys by (Stephenson, *ibid.*, Table 1). It was also found that the PSQ implemented in the General Social Survey underrepresented men who worked full time. In addition, "difficult" respondents, who are doubtlessly unrepresented in any survey with incomplete response, were probably more underrepresented in the PSQ design than the full probability design (Stephenson, *ibid.*, p. 494).
16. A memorandum from the field director (R. Nepon, First Pretest—Study 4088, unpublished memorandum, NORC, January 7, 1970) notes that the refusal rate seems no higher than normal, and does not relate to the subject matter, but rather to the usual "too busy," "not interested," and "just leaving for the store." Reported responses in the first pretest for 743 attempted contacts was: 345 not at home, 178 not in sampling universe, 84 refusals, 72 completed cases, and 64 "other," which included cases of respondent illness, non-English speakers, or contacts with businesses or vacant dwellings. This first pretest was conducted in 15 areas (Birmingham; Cleveland; Cumberland County, ME; Indianapolis; Pittsburgh; Chicago; San Francisco; Tampa; Madison County, NE; Denver; Carbon County, PA; Jasper County, SC; Hill County, MO; Carroll County, IN; and Major County, OK). Problems were encountered in Indianapolis which the field supervisor attributed to a variety of staffing problems; the "fact that Indianapolis is a conservative community and very much embroiled at present in a fight about sex education"; and "community attitudes toward Indiana University." In a second pretest, a breakdown of 2,129 attempted contacts is reported as: 926 contacts were not at home; 519 involved persons outside the sampling frame; 375 were refusals; 199 yielded completed cases; and the remaining 110 were classified as "other" (R. Nepon, Second Pretest—Study 4088, unpublished memorandum, NORC, July 13, 1970). Of the 375 refusals, 276 are described as "regular refusals" and 99 were described as study-related (including "reluctant to discuss" sex, 34; and "too personal," 17). In considering these pretest refusal rates it is important to remember that (i) pretests typically are not conducted under pressure to limit nonresponse, and (ii) the PSQ sample design does not use call-backs or encourage refusal conversion. The foregoing pretest results can be compared to the results obtained in a 1975 NORC survey of leisure time activities (NORC Survey 5059).

In the full survey, this PSQ survey design involved 7,928 potential contacts which yielded 1,136 completed interviews, 1,219 refusals, 3,400 not-at-homes, 1,684 contacts with persons outside the sampling frame, 576 cases classified as "other," and 37 cases in which contact records were missing (T. W. Smith and P. Bova, personal communications). The effects of PSQ versus full-probability sampling have been experimentally investigated (*15*).

17. Reported homosexual contact by black respondents, however, is generally close to, but slightly below that for whites, implying that any adjustment of the estimates for the distribution by race would have trivial effects.
18. National Opinion Research Center, "Specifications for survey 4088" (unpublished memorandum, University of Chicago, September 1970), p. 7.
19. In its instructions to interviewers, NORC (*18*, p. 16) also advised its interviewers that: "You have to play a double role at this point. On the one hand, you must keep yourself occupied so that the respondent doesn't feel that you are watching him and consequently become self-conscious—on the other hand, you must be observant enough to be aware if he seems to be running into difficulty and needs your help. If the respondent cannot read or indicates for some other reason that he would prefer you to ask these questions directly, it is permissible to do so." We do not know how many respondents asked for or received such assistance.
20. See R. J. A. Little and D. B. Rubin [*Statistical Analysis with Missing Data* (Wiley, New York, 1987)] for a detailed discussion of such model-based approaches to the impretation of missing data.
21. Four logistic models were fit to estimate missing values for whether the respondent had any homosexual contact. The principal model expressed the logit, $\ln[p/(1-p)]$, with p the proportion with any experience, as the sum of a constant term, −0.40 (with standard error 0.20); coefficients −0.32 (0.17) for under age 30, 0.49 (0.26), for never married age ≥30, and −0.18 (0.18) for currently or ever married age ≥30; −0.19 (0.10) for those favoring a law against prostitution and 0.19 (0.10) for those opposed; 0.30 (0.08) for those characterizing themselves as more approving than their friends of pre-marital and extra-marital sexual behavior and −0.30 (0.08) for those less approving or about the same as friends; −0.65 (0.15) for those whose opinion was that physical acts between two persons of the same sex for purposes of sexual stimulation when they love each other is always wrong, −0.07 (0.17) for those answering "almost always wrong," 0.24 (0.19) for those answering "wrong only sometimes," and 0.48 (0.14) for "not wrong at all"; and 0.55 (0.12) for those answering "yes" on the self-administered questionnaire the ability to imagine enjoyment from sexual contact with someone of the same sex, and −0.55 (0.12) for those responding "no" to this question. Because of high nonresponse to the attitude question on physical acts between persons of the same sex and to the imagination question on the self-administered questionnaire, three additional models were fit for the three remaining possibilities for presence or absence of responses to these two items, but incorporating the other independent variables in each instance.
22. For example, B. Efron, *The Jackknife, the Bootstrap, and Other Resampling Plans* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1982).
23. That is, cases that either reported no homosexual contacts or that reported some contact and also reported both the frequency of contact and the age at most recent contact.
24. We note that this estimate of homosexual experience among men who had completed college approaches that reported by Gebhard and Johnson (*6*) in their analysis of the Kinsey data after samples from prisons and homosexual organizations, such as the Mattachine Society, were eliminated from the dataset.
25. Specifically, NORC's 1988 General Social Survey [J. A. Davis and T. W. Smith, *General Social Surveys, 1972–1988: Cumulative Codebook* (National Opinion Research Center, Chicago, 1988)]. The 1988 survey employed a full probability sample design, and Michael *et al.* (*7*) report that the survey obtained a 77.3 percent response rate. Of the 638 male respondents in the survey, 69 did not complete questions asking about the number and gender of sexual partners in the last year.
26. Generally, for the men who reported their age of last contact as 20 years or older and who responded to the question on frequency-of-contact question, a rate for recent contact (same age or within 1 year of age) was derived for each combination of marital status and frequency of contact. When the observed counts were too small to form rates, some collapsing of cells across levels of frequency of contact was employed. These rates were then applied to the estimated totals, including imputations, for men with contact at age 20 years or older by frequency of contact and marital status.
27. In the 1970 Kinsey survey, currently married men account for 64 percent (SE, 12 percent) of the estimated number of men with last contact at their current age. Without adjustments for missing data, this proportion drops to 50 percent (SE, 15 percent). After examining the reported attitudes toward homosexuality and responses to the self-administered questionnaire for the six currently married men in question, T. W. Smith (unpublished memorandum) concluded that some or all of these instances arose when he respondent mistakenly or carelessly reported the sex of his sole partner (actually his wife) as male.
28. For these results, see C. F. Turner *et al.* in (*8*, appendix B). The survey after asking, "How old were you the first time you had sexual activity with someone of the opposite sex, when either you or your partner came to a sexual climax? (If the first time was when you got married, please give your age at that time.)," the age of the first partner of the opposite sex, and how often this experience occurred, the questionnaire asked, "With about how many persons altogether did you have this sexual experience before you were married? (If it happened with your husband or wife before you were first married, this counts as one person, too.)"
29. R. E. Fay, *J. Am. Stat. Assoc.* **81**, 354 (1986).
30. See, for example, the discussion of sex in prison in Kinsey *et al.* (*2*) and J. H. Gagnon and W. Simon [*Sexual Conduct* (Aldone, Chicago, 1973)]. In the 1970 Kinsey survey, 18.8 percent of the 570 men who reported military service also reported one or more same gender sexual contacts versus 12.3 percent of men who had not served in the military. (These rates do not include adjustments for missing data.)
31. The reported proportions from the 1970 census were derived from the U.S. Bureau of the Census [Census of Population: 1970, Characteristics of the Population—United States Summary (U.S. Government Printing Office, Washington, DC, 1973), vol. 1, part 1, pp. 265, 269, 275, 368, 640, and 656; Census of Population: 1970, Subject Reports, Final Report PC(2)-4E: Persons in Institutions and Other Group Quarters (Washington, DC, 1973), pp. 27, 32, 510, and 514].
32. The authors are grateful to J. Reinisch and the trustees of the Kinsey Institute for making the 1970 survey data available to us. The authors are grateful to P. Bova, T. Smith, R. Michael, and the National Opinion Research Center (University of Chicago) for providing materials relating to the execution of the 1970 survey and the sexual behavior supplement to the 1988 NORC General Social Survey. S. Gunjal assisted the authors in the data analysis. Support for this research was provided, in part, by the Rockefeller Foundation, The Russell Sage Foundation, and the U.S. Public Health Service.