DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
Centers for Disease Control

Memorandum

Date August 11, 1981

From Coordinator, KS & OI Task Force

Subject Case control study issues and hypotheses

To Task Force on Kaposi's Sarcoma and Opportunistic Infections

A. At the "Brainstorming" session on July 30, 1981, the following projects were discussed:

1. Case-control study of KS and OI involving interviews and laboratory studies.

2. Study of the immunologic status of gay men without KS and OI (e.g., in VD clinics) beginning with skin testing.

3. Descriptive epidemiology study of KS and OI through intense telephone followup of case reports.

4. Pathologic review of KS diagnostic specimens.

5. Establishment of a comprehensive national surveillance system(s) for KS and OI.

6. Toxicologic studies of inhalants.

7. Clinic-laboratory correlations of infections with *Entamoeba histolytica* (and other enteric pathogens?) among gay men, especially with respect to changes in immunologic status.

8. Prospective study of natural history of primary infection with CMV among gay men.

9. Multiple antigen serologic surveys of control groups of gay men from cities involved in hepatitis B vaccine trials (New York City, San Francisco, Los Angeles, Denver, St. Louis, Chicago).

B. There was general agreement that the design of a case/control study was the current highest priority for the Task Force. Below is an outline of some of the issues we must face in designing and implementing such a study. In most cases, I have suggested at least one solution; please criticize freely, thoughtfully, and quickly.

1. Hypotheses (study variables): The precise listing of hypotheses and study (independent) variables of interest is critical to our case definition, selection of cases and controls, sample size determinations, etc. The attachment is a partial list of hypotheses and study variables. Please add to those or criticize them in writing. The session on Thursday, August 13, at 10:00 a.m. will specifically address this issue.

2. Case Definition and Selection of Cases: Cases should be selected to maximize the possibility of detecting differences in causal variables between them and controls - i.e., uncertain or unrelated cases should not be included. If sufficient numbers of cases were available, it would be preferable to include only incident cases, since prevalent cases will tend to be more ill, have poorer recall, be more influenced by past interviews or publicity, and be less valuable for some laboratory studies. Since many of the study variables of interest involve patient recall, living cases are preferable to dead cases. (Some variables of potential interest could be ascertained by interviewing relatives or reviewing charts, but controls would probably need to consist of similar age homosexual men dying of unrelated disorders for these cases.) To date, we know of approximately 120 cases of KS and/or OI. Approximately 80 are living homosexual men. New cases are being reported to us at the rate of 5-10 per week, but a number of these are dying or have died at the time of the report.

   Suggested Case Definition and Selection: Living homosexual men with KS and/or OI without serious underlying disease who are not in serious or critical condition medically. Both incident and prevalent cases should be included and the results compared for comparability.

3. Selection of Controls: This is perhaps the most important consideration in case/control studies. Ideally, controls would be similar to cases in all ways except exposure to the causal variable(s). Matching of cases to controls (or stratification in analysis) is used to eliminate differences between cases and controls which do not reflect differences in the variables of primary interest (presumably the causal variables). Matching for these "confounding variables" is more efficient (than stratification) when the variable is uncommon in control populations to be studied (e.g., homosexuality and age), "Intangible" confounding factors are often dealt with by selection of controls from the same area and/or SES as cases (e.g., city). In interview case/control studies, hospital controls are often used to reduce expense and facilitate comparability of the obtaining of information with hospitalized cases. (Ideally, all cases and controls would be interviewed in the same place by the same person who was blinded to their status. Alternatively, the effects of interviewing healthy controls, say at home or in a crowded VD clinic, versus ill cases in a hospital room need to be considered.) Generally, sources of controls should be other than those influenced (positively or negatively) by the factors under study. This would suggest that VD clinics and religious organizations would not provide optimal controls for the study of drug usage and sexual behavior. Finally, selection of controls should be done systematically and differences in response rates between cases and controls noted. (For laboratory studies, for example, the type of person "volunteering" to donate 80 cc of whole blood may be different from a "typical" control.) When incident cases of an uncommon disease are studied, controls should probably be matched individually to cases as they occur to assure comparability in time of interview, etc. While matching cases with controls can result in statistical efficiency, the effects of the variables matched cannot be evaluated (e.g., sex preference, race, age, and city). Also, overmatching can occur when variables intermediate between the disease and the study variables are matched (e.g., CMV infection).

Since controls will be much more plentiful (hence, cheaper) than cases in this study, multiple controls from different sources could be selected for each case.

Suggested Selection of Controls: Four controls for each case should be individually matched by sex, sex preference, race, and age (within $\pm$ 3 years). Two controls each would be selected for each case from two sources in the same city (e.g., VD clinics and a systematic sampling of men from a gay community organization). Cases and controls would be interviewed by the same individual in the same manner. Acceptance as a control for interview would not be contingent upon willingness to provide laboratory specimens.

4. Comparability of Information Collected on Study Variables Between Cases and Controls: Ideally, a single blinded unbiased professional interviewer would obtain information on all cases and controls in the same setting. None of the cases (or controls) would have been previously interviewed and they would have been unaffected or equally affected by loss of memory, guilt, publicity, etc.

Comparability can be assured to a lesser extent by having the same person interview cases and controls in a standard manner. By comparing the time spent in interviews, the estimates of the reliability of the respondents, and answers to "dummy variables" (e.g., place of birth, history of chickenpox) between cases and controls, comparability can be assessed retrospectively. Some responses can be validated for both cases and controls by checking available records which are independent of the study (e.g., syphilis registry).

Suggested Alternative: EIS Officers or other physicians at CDC personally interview cases and matched controls in a standard manner. Cases and their matched controls should be interviewed by the same individual and times and estimates of reliability recorded.

Jim Curran

James W. Curran, M.D.

ATTACHMENT

KAPOSI'S SARCOMA AND OPPORTUNISTIC INFECTIONS TASK FORCE
HYPOTHESES, STUDY VARIABLES, AND "CLUES"

A. The likely strong association of these disorders with active homosexual men suggests environmental and/or behavioral factors as the primary cause(s). (Though genetic factors may be operative in determining susceptibility among homosexual men, they are unlikely to explain either the increase in occurrence of these disorders or their predominance among homosexual men.) It seems that, at this point, this association should direct our hypotheses.

B. Clues:

1. From available data:

   a. Cases have been concentrated in New York City and California (undoubtedly there is a bias in reporting, but what hypotheses would this geographic concentration suggest if it is real?) Most cases were not born in New York City or California.

   b. The ages of the men interviewed are older than very sexually active gay men seen in public STD clinics or responding in the Jay-Young Survey (38 = mean, 35 = median).

   c. The men seem to report having a larger number of different contacts than expected.

   d. Most of the men reported use of inhalants frequently.

   e. The range of years of practice of homosexual activity was 6-37 years (median 16). These men are not "new" to the homosexual lifestyle, nor do they seem to currently lead "dual" lives.

   f. The men have a median of 16 years of education; most are employed. They do not seem to be lower SES.

2. Among other things, these variables characterize sexually active homosexual men:

   a. High incidence of syphilis, gonorrhea, condyloma acuminata, genital herpes, and other sexually transmitted diseases.

   b. Very high incidence of hepatitis (B, A, non-A-non-B, ?CMV, etc.).

   c. Use of inhalants in greater frequency than heterosexuals.

   d. Higher frequency of active and passive oral-genital, genital-rectal, and oral-g~~enital~~ rectal sexual activity.

e. Much higher frequency of enteric infections (parasitic, bacterial, viral) than heterosexuals.

f. A tendency to associate with each other (and to exclude others) in clubs, baths, bars, etc.

g. ?Use of rectal lubricants, douching equipment, enemas, etc.

3. What is suggested by related cases?

   a. Which diseases do they have?

   b. What was nature of the relationships?

   c. What was relation of exposure periods to disease prodrome or onset? (?incubation period)

4. What is suggested by cases occurring in women and heterosexual men?

   a. Why no KS? (or only one case)

   b. How does PCP differ? (?prodrome, immunologic status, etc.)

5. Why is KS occurring to the apparent exclusion of other malignancies (compared, for example, to renal transplant populations).

6. Why is PCP accounting for the majority of opportunistic infections reported?

7. What clues can we get from the literature?

   a. Male predominance in KS.

   b. Lesser KS male predominance among younger men (at least in Africa).

   c. Much higher KS incidence and different age distribution in Africa.

   d. Decreasing KS severity with increasing age (at least in Africa).

   e. Higher KS incidence among renal transplant recipients.

   f. Spontaneous remission of KS when immunosuppressants are discontinued in renal transplant patients. (What precise immunologic changes occur in this situation? Are there any naturally-occurring similar situations?)

   g. What is precise pathogenesis of PCP among drug addicts?

   h. What is precise pathogenesis of PCP among infants without underlying disease?

   i. What about the report of 3 cases of PCP in a family?

j. What is precise pathogenesis of PCP in patients with protein-calorie malnutrition (Kwashiorkor)? What immunologic changes occur in this disease?

k. What is nature and cause of immunologic dysfunction in heroin addicts? (?due to chronic hepatitis as suggested.)

l. What is meaning of association of HLA B8/DRW3 hablotype with immunologic dysfunction?

m. etc.

C. Hypotheses:

1. An environmental agent (microbial or chemical) is causing lasting immunosuppression among homosexual men. As an intermediate variable, immunosuppression is rendering some of these men susceptible to OI's and allowing the expression of KS.

2. Repeated CMV infection or reactivation is occurring resulting in immunosuppression, etc. This "continuous antigenic stimulation" may lead to an increased synthesis of defective viruses. These defective viruses may be oncogenic in certain individuals.

3. A heavy total body infection burden is exhausting the individual's ability to respond, resulting in immunosuppression, etc.

D. Study Variables

1. Behavioral

a. sexual activity

b. nonmedical drug usage

c. interactions of a and b.

d. other social variables.

2. Demographic

a. ethnicity, race

b. residence (duration)

c. place of birth

d. travel history

e. present and past marital status

f. SES, income

g. family history

h. education

3. Occupational

    a. occupation(s) (duration)

    b. training

    c. solvents, chemicals, radioactive materials

    d. income

4. Past medical history

    a. prodromal illness (duration, characteristics)

    b. STD's

    c. hepatitis (type, time, response)

    d. mononucleosis (type, time, response)

    e. tonsillectomy (time, etc.)

    f. GI disturbances and diseases

    g. medication history

    h. dietary history (changes, etc.)

    i. alcohol, tobacco

    j. childhood illnesses

    k. friends or partners with mono, etc., persistent fever, lymphadenopathy

    l. friends or partners with unexplained deaths, cancer, etc.

DEPARTMENT OF HEALTH & HUMAN SERVICES

File

Public Health Service
Centers for Disease Control

Memorandum

Date August 17, 1981

From Keewhan Choi, Mathematical Statistician
EPO, Consolidated Surveillance and Communications Activity

Subject Case/Control Study

To Dr. James W. Curran
Coordinator, Task Forces on Kaposi's Sarcoma and Opportunistic Infections

After having discussed the case/control study issues with you and Steve Thacker, I have come up with the following recommendations:

1. Selection of cases. Not only the living incident cases, but also dead cases should be selected.

2. Selection of Controls. For each gay male case, four controls should consist of two gay males, one straight male, and one straight female. Cases and controls should be matched by race and city only, but not by age. One of the two gay controls may be selected from VD clinics, but others should be from gay community organizations. (We may need to revise our selection criteria for controls when objectives of the case/control study are firmly agreed upon by the task force).

The aforementioned recommendations for the selection of subjects are based on the assumption that the planned study is exploratory in nature rather than confirmatory. In an exploratory case/control study, we have a long list of variables to be examined and the main objective of the study is to identify a set of important (potential) risk factors. In such a study, it is essential to have as many varied control groups as possible, because the analytical method for case/control studies boils down to searching for subtle differences between cases and controls, often hidden in many correlated variables.

Inclusion of the dead cases into the study is important because comparison of important (potential) risk factors for the dead cases with those for the living cases will provide us with a test for an internal validity as well as a possibility for identification of additional risk factors for the fatal cases.

Since the effects of matched variables on the disease cannot be evaluated, no variable should be matched unless we are certain that it is not a risk factor. Hence, cases and controls should not be matched by age for the case/control study.

Keewhan Choi

Keewhan Choi, Ph.D.

FE

DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
Centers for Disease Control

Memorandum

Date September 11, 1981

From Mathematical Statistician
Consolidated Surveillance and Communications Activity, EPO

Subject Kaposi's Sarcoma Case-Control Study Protocol

To Dr. James Curran
Coordinator, KSOI
Through: Chief, CSCA, EPO SBT

Steve Thacker and I have studied the KS case-control study proposal and would like to make the following two suggestions:

1. Field test the questionnaire to a dozen or so people (VD clinical patients) in Atlanta to see if usual answers to the questions can be obtained. With the field test results, we can begin immediately to work on the coding of data and computer programs to analyze the data. (We need the lead time so that we will be able to analyze the complete case-control study data as soon as the data are cleaned up after the interviews are completed.

2. In each source of the controls, select controls for the study <u>RANDOMLY</u>, <u>not</u> the first person who meets the selection criteria. For example, prepare a list of two digit random numbers for each clinic and choose the kth patient of the day if k is the number in the list.

   Random selection of subject is an absolutely essential ingredient in the case-control study, which enables us to infer the study results to a population beyond the study subjects.

Specific comments on the questionnaire are made in the attached copy of the protocol. We presume Walter Gunn is working on questionnaire format and design.

Keewhan Choi
Keewhan Choi, Ph.D.

Attachment

cc:
Dr. Harold W. Jaffe