AMERICAN JOURNAL OF EPIDEMIOLOGY
Vol. 116, No. 6

THE EFFICACY OF A POPULATION-BASED COMPARISON GROUP IN CROSS-SECTIONAL OCCUPATIONAL HEALTH STUDIES

PAUL A. SCHULTE, MITCHELL SINGAL, WILLIAM T. STRINGER, JOHN R. KOMINSKY AND PHILIP J. LANDRIGAN

Schulte, P. A. (NIOSH, Cincinnati, OH 45226), M. Singal, W. T. Stringer, J. R. Kominsky and P. J. Landrigan. The efficacy of a population-based comparison group in cross-sectional occupational health studies. *Am J Epidemiol* 1982;116:981-9.

The availability and the choice of appropriate comparison groups are essential for valid occupational epidemiologic studies. Too often, however, adequate comparison groups cannot easily be found within a workplace environment or extracted from the general population. An evaluation of the efficacy of using a pool of comparison subjects from the Health and Nutrition Examination Survey (HANES) was performed on data gathered by the National Institute for Occupational Safety and Health in 1979. Comparison groups from the HANES pool were derived for 246 workers at four different commercial/industrial facilities in the Niagara Falls, New York, area and the comparability between the groups was assessed for several demographic, behavioral, and biomedical variables. The HANES groups exhibited a high degree of comparability with regard to most variables, excluding ancestry. The HANES pool may serve as a useful source of subjects to allow for the comparison of disease rates where occupational exposure is the key distinguishing feature between groups.

cross-sectional studies; epidemiologic methods; occupational diseases

The choice of comparison groups is one of the most vulnerable aspects of epidemiologic studies, particularly occupational health studies (1–4). Depending on which groups are used for comparison, a true excess prevalence of a disease may be masked or exaggerated (2). Conversely, a spurious excess of disease may appear because of a lower background prevalence in a comparison group chosen without regard to selection bias. The need for adequate comparison groups is evident in the increasing number of cross-sectional epidemiologic studies performed for preliminary evaluations of workplace environments (5).

In the workplace environment, it is often difficult to find adequate unexposed groups suitable for comparison with the study group. This is true for a number of reasons: 1) a workforce may not be large enough to allow for the identification of both an exposed and unexposed group with regard to a given chemical or process; 2) intra-workplace job changes may result in a majority of the workforce with a history of employment in the target

Received for publication January 25, 1982, and in final form May 19, 1982.

Abbreviation: HANES, Health and Nutrition Examination Survey.

From the Division of Surveillance, Hazard Evaluations, and Field Studies, National Institute for Occupational Safety and Health, Cincinnati, OH. (Send reprint requests to P. A. Schulte, Mail Stop F-10, NIOSH, 4676 Columbia Parkway, Cincinnati, OH 45226.)

The authors thank Robert Murphy of the National Center for Health Statistics for his assistance.

Information, publications of HANES data, and a catalog of published data tapes are available from the Scientific and Technical Information Branch, National Center for Health Statistics, Center Building, Room 1-57, 3700 East-West Highway, Hyattsville, MD 20782.

area; and 3) differential demographic characteristics may exist between those workers in the study group and those in the rest of the plant, so that meaningful comparisons are obscured by confounding factors (6). The choice of a comparison group from outside the work environment also presents a problem, since the use of general population groups may lead to a comparison bias known as the "healthy worker effect" (2, 7). This phenomenon is frequently observed in occupational health studies, presumably because some degree of good health is a prerequisite to employment.

Cross-sectional occupational health studies, while limited in their utility, provide the opportunity for rapid determination of diseases with either a high incidence or long duration (8). Such studies may also detect chronic, nondebilitating diseases such as those with a recovery potential that will allow re-entry into the workplace. Cross-sectional data may be gathered not only on diseases or conditions prevalent at a given time, but also on those experienced in the past. The gathering of such data allows for the evaluation of the current health of a workforce or, when repeated on the same workforce, an appraisal of historic trends. Both of these evaluation efforts depend on which comparison group is used.

The need for adequate comparison groups and the problems and costs involved in finding such groups for each study are ample reasons to search for currently available pools of comparison subjects. The Health and Nutrition Examination Survey (HANES) of the National Center for Health Statistics provides a multistage, stratified probability sample of the civilian noninstitutionalized population of the United States (9, 10). The survey provides data on the prevalence of specific conditions and chronic diseases, as well as normative data on such requirements as physical and biochemical parameters and information on nutritional status and deficiencies (11). Over 40 collaborative epidemiologic projects have been initiated using HANES data, but few have involved investigation of occupational disease (12). The efficacy of using this group as a comparison standard depends on two factors: 1) whether such a sample or its subsets are by design or adjustment similar to target groups in the distribution of various potentially confounding factors, and 2) the existence of no other significant, undetected sources of bias within the sample which could lead to unreliable conclusions. This paper evaluates the use of comparison groups drawn from the HANES sample in exposure-selective cross-sectional studies of four commercial/industrial facilities in the same community.

Materials and methods

Occupational groups from four commercial/industrial facilities in the Niagara Falls, New York, area were chosen for comparison with the HANES group. They included employees in a steel finishing company (Group A), workers in two companies that manufactured shipping drums and metal alloy powders, respectively (Groups B and C), and the maintenance and housekeeping workers of a university (Group D). All were being studied because of their workplaces' proximity to a landfill containing toxic chemicals (13). Since they all were potentially exposed to vapors and dusts from the landfill, comparison groups from nonexposed areas had to be chosen and their comparability evaluated. The HANES I Survey (1971–1973) of 6913 adults, ages 25 to 74 years, provided a pool of comparison subjects (8). Persons in the HANES sample had been administered a variety of questionnaires and given physical examinations and laboratory tests at the time of the HANES survey.

The four occupational study groups were given identical questionnaires containing selected sections of the HANES questionnaire. These questionnaires were administered by a team of interviewers trained for the purpose. The results of these questionnaires were compared with results obtained from a subset of the HANES sample comprised only of people who indicated that they were employed at the time of the survey. Two comparison subjects were chosen for each subject in each study group. To achieve maximum comparability, the members of each study group were individually matched with two members of the HANES group on age (±3 years), sex (same), race (study subjects were classified as "black" or "white," the latter including, for matching purposes, all non-blacks), marital status (married, never married, formerly married), and household income (high, medium, and low, adjusted for inflation). Consequently, groups A, B, C, and D had 65, 60, 97, and 21 matched triples, respectively. The comparability between the study groups and the HANES matches was evaluated for demographic variables such as schooling and ancestry, for behavioral variables such as smoking and drinking habits, and for selected health and biologic variables such as a history of rheumatic fever, heart murmur, allergy to food, history of polio, and weight. These health variables were chosen because, with the possible exception of weight, they are presumably unrelated to any occupational exposure.

The relative similarity between each occupational group member and his or her HANES matches was determined for the variables identified using the Mantel-Haenszel method for matched triples with dichotomous outcomes, and the expression for the Stuart-Maxwell chi-square statistic for three outcomes derived by Fleiss and Everitt (14–16).

Results

A comparison of the number of years of schooling completed showed that no statistically significant difference existed for the distribution of school years between any occupational group and its matches or for all occupations combined (table 1). In 43 per cent of the matched triples, the educational levels were the same. The HANES matches had more education in 25 per cent of the triples and less in 32 per cent. Specifically, three of the four occupational groups had a higher proportion of persons with some high school education than their HANES counterparts. Two of the HANES groups had achieved a greater proportion of some college education and two had a lesser proportion than their specific occupational group.

Considerable differences existed for ancestry or national origin in each occupational group and their HANES comparison groups (table 2). The occupational groups were all comprised of significantly more people of Italian or Polish ancestry, and fewer of Irish ancestry; the occupational groups had from 9 to 38 per cent Italian members, while the percentage of Italians in their matches ranged from 0 to 8.4 per cent. The occupational groups also had practically no Spanish, Mexican, Chinese, or American Indian members. As was expected because of matching for race, the groups were almost identical for members of "Negro" origin.

Table 3 shows the per cent who were current cigarette and cigar smokers for each occupational group and its comparison. Three of the four groups were not significantly different from their HANES comparisons for cigarette smoking ($p < 0.05$). The percentage of current cigarette smokers ranged from 53.4 to 81.0 per cent in the occupational groups and 56.1 to 61.7 per cent in the comparison groups. The least "industrial" workers, that is, the maintenance/housekeeping workers of the university, had the highest smoking rate, 81.0 per cent. Conversely, they had the lowest cigar smoking rate, 13.3 per cent. Three individual study groups had higher prevalences of cigar smoking

TABLE 1

Percentage distribution of years of education completed by occupational groups A, B, C, and D in Niagara Falls, New York, and their HANES comparisons, 1979*

Occupational groups and comparisons		Years of education completed††		
		Elementary school 1–8	High school 9–12	College 13–17
A	n = 65	13.7	64.6	21.5
HANES (A)	n = 130	18.0	49.0	32.8
	χ^2 = 1.33			
B	n = 60	20.4	64.5	14.3
HANES (B)	n = 120	30.9	52.1	17.2
	χ^2 = 1.96			
C	n = 97	10.3	57.7	32.0
HANES (C)	n = 194	20.2	55.9	27.0
	χ^2 = 4.6			
D	n = 21	23.9	47.6	28.6
HANES (D)	n = 42	22.5	60.0	17.5
	χ^2 = 0.071			
A,B,C,D‡	n = 246	14.6	60.8	24.4
HANES (A,B,C,D)	n = 492	22.0	52.6	25.0
	χ^2 = 1.53			

* Health and Nutrition Examination Survey.

† Matched triple analysis for categories 1–8, 9–12, 13–17, χ^2 with 2 d.f.

‡ Combined analysis includes three triples from a fifth occupational group not in the individual analyses.

than their HANES comparisons, an association that was statistically significant ($p < 0.05$) for the study population as a whole.

Alcohol consumption by proportion and magnitude did not differ significantly between three of the occupational study groups and their comparison groups but did in group C (table 4). When all four groups were taken together and compared with pooled HANES groups, no statistically significant difference was seen. Generally, however, the study groups had lower percentages than comparison HANES groups for the designation "everyday drinking," with percentages ranging from 9.5 to 14.8 per cent, compared with those of the comparison group of 15.9 to 19.3 per cent. With each category of lower frequency of consumption, the occupational groups and their HANES matches became more similar in terms of drinking, and the study group exceeded the comparison groups in the category of drinking "two to three times per week."

Several biologic markers were reviewed to evaluate the comparability between groups. No difference in weight was found between study groups and their HANES comparison groups. The prevalences of disease conditions such as a history of rheumatic fever, heart murmur, allergy to food, and history of polio were evaluated and no statistically significant differences between groups were found. Two other biologic variables which would have been considered unrelated to chemical exposure were hiatus hernia and surgery for hiatus hernia. Two of the four study groups, however, had significantly greater frequencies than their comparisons for these two conditions ($p < 0.05$).

DISCUSSION

Generally, the study groups and their matches were similar in regard to most of the behavioral, demographic, and biologic variables evaluated. While they were racially the same, due to matching, they did show differences in reported ancestry,

TABLE 2

Percentage distribution of ancestry in occupational groups A, B, C, and D in Niagara Falls, New York, and their HANES† comparisons, 1979*

Ancestry or origin	Occupational groups and comparisons							
	A	HANES (A)	B	HANES (B)	C	HANES (C)	D	HANES (D)
Irish	9.2	14.8	13.3	23.7	7.2	13.1	14.3	15.0
Italian	9.2	0.0	38.3	2.5	24.7	8.4	19.0	0
French	9.2	8.6	8.3	1.7	1.0	6.3	0	5.0
Polish	12.3	2.3	10.0	0	10.3	2.6	4.8	5.0
Russian	0	0	0	0	2.1	1.0	0	5.0
English	15.4	19.5	6.7	14.4	14.4	13.1	19.0	20.0
Spanish	0	2.3	0	1.7	0	1.6	0	2.5
Mexican	0	2.3	0	4.2	0	3.7	0	5
Chinese	0	1.6	0	1.7	0	0	0	0
Amindian	0	2.3	0	1.7	0	2.1	0	0
Negro	3.1	3.1	5.0	5.1	8.2	8.4	0	0
American	0	7.0	0	11.0	0	6.3	0	11.0
Jewish	0	0	0	0	0	0	0	2.5
Other	29.2	12.5	8.3	11.9	16.5	14.7	23.8	15.0
N =	65	128	60	118	97	191	21	40

* Per cents do not add to 100.0% because all ancestral groups are not represented.
† Health and Nutrition Examination Survey.

TABLE 3

Percentage distribution of cigarette and cigar smoking in occupational groups A, B, C, and D in Niagara Falls, New York, and their HANES comparisons, 1979*

Occupational groups and comparisons	Now smoke cigarettes		Now smoke cigars	
A	67.3	$\chi^2 = 1.28$	23.1	$\chi^2 = 1.33$
HANES (A)	61.7		35.1	
B	81.0	$\chi^2 = 6.20$†	13.3	$\chi^2 = 2.33$
HANES (B)	56.1		41.4	
C	53.4	$\chi^2 = 0.08$	20.7	$\chi^2 = 0.48$
HANES (C)	59.1		37.7	
D	64.7	$\chi^2 = 0.11$	50.0	$\chi^2 = 1.00$
HANES (D)	57.1		30.0	
(A,B,C,D)	64.0	$\chi^2 = 3.76$	21.9	$\chi^2 = 5.54$†
HANES (A,B,C,D)	59.0		37.2	

* Health and Nutrition Examination Survey.
† $p < 0.05$.

presumably reflecting differences between local and national ethnic composition. The study groups tended to have lower proportions of cigar smokers, but (with one exception) tended to have similar proportions of cigarette smokers as their HANES matches. Cigar smoking, however, is not known to have the extensive effects on health that cigarette smoking does. On the basis of this evaluation, and despite the differences in ancestry, the authors concluded that the comparison groups chosen from the HANES sample were adequate for use as referents in the cross-sectional medical study conducted by NIOSH at the landfill in Niagara Falls, New York (13).

This study has shown that matched

TABLE 4

Percentage distribution of alcohol consumption in occupational groups A, B, C, and D in Niagara Falls, New York, and their HANES comparisons, 1979*

Frequency of consumption	Occupational groups and comparisons									
	A	HANES (A)	B	HANES (B)	C	HANES (C)	D	HANES (D)	A,B,C,D	HANES (A,B,C,D)
Every day	11.5	19.3	14.8	15.9	9.8	17.4	9.5	18.8	11.3	17.8
Just about every day	9.8	7.9	11.1	11.0	6.5	7.8	14.3	9.4	9.1	8.5
About 2 or 3 times/week	24.6	12.3	31.5	18.3	35.9	22.2	19.0	18.8	29.9	18.3
About 1–4 times/month	31.1	34.2	27.8	30.5	28.3	31.7	33.3	31.3	29.4	32.0
About 12/year	13.1	16.7	5.6	11.0	10.9	8.9	9.5	0	10.0	11.0
3 times/year	9.8	9.6	9.3	13.4	8.7	12.6	14.3	21.9	10.4	12.5
χ^2†	2.96		3.70		10.98‡		0.53		2.49	

* Health and Nutrition Examination Survey.
† Chi-square with 2 d.f. for frequency categories 1 + 2, 3, 4 + 5 + 6.
‡ $p < 0.005$.

comparison groups derived from the HANES national sample exhibit an acceptable degree of comparability with selected occupational study groups for demographic, behavioral, and biologic factors that could potentially confound associations between occupational exposures and their effects. Whether this comparability is adequate to minimize confounding by those and related factors, cannot yet be determined with complete assurance, since only a few variables so far have been studied.

The efficacy of the HANES pool as the source of comparison groups in occupational health studies results from two major features. First, the HANES pool was large enough to allow for the selection of an employed subset. The HANES pool was chosen in such a fashion as to avoid unintentional selection or sampling biases and to minimize non-response. Further, it is relatively inexpensive to use the HANES pool as a source of comparison subjects. Usually, population comparison groups are expensive to obtain if they need to be developed by sampling and interviewing *de novo*. This effort and expense does not have to be borne by investigators utilizing HANES.

The second attractive feature of the HANES pool is the capacity, when appropriate, to employ extensive matching as part of the study design. Each study subject had two matched counterparts, and the matching allowed for exact and relatively narrow categories. It was not surprising that, as a result of matching, other related factors such as years of schooling and smoking and drinking habits were similar. The finding of similarity in body weight is added confirmation of the comparability of the groups.

The quality of the information gathered was considered relatively equal since both the HANES group and the study groups were asked the same questions in much the same way. The external validity of these data has not been fully ascertained, but Melius (17) found that the HANES employed group appeared to be comparable in terms of many demographic variables with the US population at the time of the survey. Subsets of the HANES pool are quite likely to have variable disease rates compared with the whole. This should not affect the internal validity of the study because of the matching.

Less reassuring was the failure of the matched group members to have the same ancestry or national origin as the study subjects. This discrepancy obviously reflects differences between the overall US ethnic composition and that of the study locality. This difference may have minimal influence in some studies but could be a source of bias in others, if a study, for example, involves variables that are highly culturally conditioned, such as diet, smoking, or other personal customs, habits, or recreational patterns (18). The HANES questionnaire includes questions on diet which, although not utilized in this study, could be used in others.

Other limitations of this study are also of concern. Some of the questions chosen from the HANES survey are not sufficiently specific to allow detailed evaluation of comparability. For example, the question on smoking, "Do you smoke cigarettes now?", does not indicate the amount. There is a question in HANES on amount, but it was not used in this study. Also, the HANES questionnaire makes no allowance for type of cigarette, such as filter or nonfilter, factors which could significantly affect the "dose" of smoke constituents. Variable participation rates also may have influenced aspects of this study. Participation by the four study groups ranged from 42 per cent to 79 per cent. While low participation rates are a potential weakness of cross-sectional studies, in this particular study the problem was not of great importance as far as the comparability between respondents and their HANES matches was concerned. It is not known, however, if the non-participants differed in any epidemiologically important way from the participants.

Another inherent limitation of the application of the HANES data base is the need to use the same questions in a particular study as were used in the HANES survey. It is possible to use abbreviated versions of the HANES, but we suspect that the wording of the individual questions and their relationship to associated questions must be kept the same for data comparability.

The HANES sample was chosen from the years 1971–1973, and thus the possibility of some temporal bias arises. In the study reported here, it was necessary to utilize an inflation adjustment factor of 1.6 (derived from the Consumer Price Index) to adjust the income levels prior to matching. Both the study groups and the HANES matches were asked to report on their prior and current health status. Since these interviews were performed at different times, by as much as nine years, it may be that societal trends or specific events may allow for significant recall bias to be introduced in the response of the study groups. The potential for this recall bias could be increased by some precursor catastrophe, outbreak, or revelation that prompted the study. This could result in inflated response rates. Some of the temporal and recall biases may be reduced for future users of the HANES pool. Starting in late 1981 the National Center for Health Statistics began to go back to the people surveyed between 1971–1975. In addition to prevalence data, this new survey will allow for estimates of the incidence of selected conditions as well as for the study of slow-acting consequences of long term and low-dosage exposure to a combination of environmental, dietary, social and demographic factors (19).

The HANES data base may be used in various ways to provide comparison information for epidemiologic studies. The most common approach has been the application of age-sex-specific rates to a study group to determine the expected or standardized rates for comparisons with observed rates (12). Another approach, used in this study, is matching.

Matching is the use of constraints in the selection of the comparison groups which are aimed at making the distribution of the potentially confounding variables within comparison groups as similar as possible to the corresponding distribution in the study group (20). Matching is useful where there is a relatively large pool of potential comparison subjects (21), as in the case of the HANES data base. In this instance matching also is cost-efficient, as there is a large number of potential matches for each study subject. The accuracy of matching in exposure-selective cross-sectional studies meets two of the criteria of Billewicz (22), namely that the outcome variables are generally dichotomous and the study group sample size is small. The third criterion, that matching variables be highly correlated with outcome illnesses, is dependent on the specific outcomes being investigated. It is possible to question whether there exists in this study the possibility of over-matching, that is, that one or more of the matched variables may be part of the causal pathway between the exposure and disease under study, and that matching on such variables will affect the association (22, 23). Monson (2) and Anderson et al. (21) have provided criteria to consider when designing studies involving matching.

Matching in exposure-selective cross-sectional studies often leads to increases in size efficiency, especially if stratified analysis is performed using strata of the matching factors (20). In this study, the size efficiency was increased by matching because the factors used in the matching have been shown to be highly correlated with many of the diseases being investigated. In matched-design analyses, it is also possible to retain the matching and still control for additional confounding variables by methods previously described (21, 24, 25).

Matching with more than one comparison subject allows for an increase in the power of the chi-square test. This is particularly relevant in studies of occupational diseases where failure to detect elevated prevalence rates may lead to continued exposure to the substances or conditions under study. Using the method of Miettinen (26), the power for the current study was determined to be 0.25 for a rate difference of 5 per cent, 0.70 for a difference of 10 per cent, and 0.95 for a difference of 15 per cent for the 246 triples. For a representative plant (Plant B) with 60 subjects and 120 matches, powers of 0.50, 0.70, and 0.90 were calculated for rate differences of 15, 20 and 25 per cent, respectively.

Presently, the HANES comparison pool has not received sufficient evaluation as a possible tool for occupational epidemiology. For example, there exists the opportunity to account for occupation by matching, adjusting, or selecting specific occupational sub-populations for comparison. Other potentially confounding variables may be handled similarly. Further efforts are needed to document the strengths and limitations of the HANES sample as a source of comparisons in cross-sectional occupational health studies. The results of this study indicate that HANES may be useful in these types of studies.

References

1. Enterline PE. The estimation of expected rates in occupational disease epidemiology. Public Health Reports 1964;79:973.
2. Monson RR. Occupational epidemiology. Boca Raton: CRC Press, Inc., 1980.
3. Hulka BS, Grimson RC, Greenberg BG, et al. Alternative controls in a case-control study of endometrial cancer. Am J Epidemiol 1980; 112:376-87.
4. Hernberg S. Epidemiologic methods in occupational health research. Work Environ Health 1974;11:59-68.
5. National Institute for Occupational Safety and Health. Annual Report to Congress. Cincinnati, OH: National Institute for Occupational Safety and Health, 1980.
6. McMichael AJ, Gerber WS, Gamble SF, et al. Chronic respiratory symptoms and job type

within the rubber industry. J Occup Med 1976;18:611-17.
7. McMichael AJ. Standardized mortality ratios and the "healthy worker effect": scratching beneath the surface. J Occup Med 1976;18:165-8.
8. Feinstein AR. Clinical biostatistics: The architecture of cross-sectional research (parts 1 and 2). Clin Pharmacol Ther 1978;23:81-96, 487-93.
9. National Center for Health Statistics. Plan and operation of the health and nutrition examination survey, United States 1971-1973. Rockville, MD: National Center for Health Statistics, 1974. (Vital and health statistics. Series 1, Number 10a) (DHEW publication no. (HRA) 76-1310).
10. National Center for Health Statistics. Plan and operation of the HANES I augmentation survey of adults 24-74 years, United States 1974-75. Rockville, MD: National Center for Health Statistics, 1978. (Vital and health statistics. Series 19, Number 14) (DHEW publication no. (HRA) 78-1B14).
11. Smith SS. Major data systems of the National Center for Health Statistics. Public Health Reports 1981;96:200-1.
12. Hattwick M. The National Health Examination Surveys Collaborative Epidemiology Program. In Abstracts of Papers Presented at the Eleventh Annual Meeting of the Society for Epidemiologic Research, Iowa City, June 14-16, 1978. Am J Epidemiol 1978;108:243.
13. National Institute for Occupational Safety and Health. Health hazard evaluation report 79-022-789. Cincinnati, OH: Centers for Disease Control, 1980.
14. Mantel N, Haenszel W. Statistical aspects of the analysis of retrospective studies of disease. JNCI 1959;22:719-48.
15. Fleiss JL, Everitt BS. Comparing the marginal total of square contingency tables. Br J Math Stat Psychol 1971;24:117-23.
16. Fleiss JL. Statistical methods for rates and proportions. New York, NY: John Wiley and Sons 1973.
17. Melius J. Dissertation on use of HANES results for cross-sectional occupational studies (in progress).
18. MacMahon B, Pugh TF. Epidemiology: principles and methods. Boston, MA: Little, Brown and Company, 1970.
19. Rice DP. New ventures in health statistics. Public Health Reports 1981;96:207-11.
20. Kupper LL, Karon SM, Kleinbaum DS, et al. Matching in epidemiological studies: validity and efficiency considerations. Biometrics 1981; 37:271-91.
21. Anderson S, Anquier A, Hauck WW, et al. Statistical methods for comparative studies. New York, NY: John Wiley and Sons, 1980.
22. Billewicz WZ. The efficiency of matched samples: an empirical investigation. Biometrics 1965;21:623-44.
23. Miettinen OS. Matching and design efficiency in retrospective studies. Am J Epidemiol 1970; 91:111-18.
24. Prentice R. Use of the logistic model in retrospective studies. Biometrics 1976;32:599-606.
25. Rosner B, Hennekens C. Analytical methods in matched pair epidemiological studies. Int J Epidemiol 1978;7:367-72.
26. Miettinen OS. Individual matching with multiple controls in the case of all-or-none responses. Biometrics 1969;25:339-55.