Bioelectromagnetics Supplement 5:S32–S47 (2001)

The Potential Impact of Bias in Studies of Residential Exposure to Magnetic Fields and Childhood Leukemia

Daniel Wartenberg*

Environmental and Occupational Health Sciences Institute, UMDNJ—Robert Wood Johnson Medical School, Piscataway, New Jersey

Bias can have a major impact on the results of epidemiologic studies. In investigations of the possible association between residential exposure to magnetic fields and the occurrence of childhood leukemia, many have raised questions about selection bias, including participation bias and information bias. In this review, the data on these possible sources of bias are summarized and their likely impact is evaluated. Most data suggest that if a bias exists, it is a bias towards the lack of association between exposure to magnetic fields and childhood leukemia. In addition, given the wide variety of study populations and measurement protocols, it is unlikely that a single design flaw has resulted in consistent effects across all studies and can be the sole explanation for the reported associations. Bioelectromagnetics Supplement 5:S32–S47, 2001. © 2001 Wiley-Liss, Inc.

Key words: selection bias; participation bias; information bias; epidemiology; study design flaws; emf; leukemia

INTRODUCTION

Bias is generally defined as the presence of systematic errors in the results of an epidemiologic study, the finding of a false effect or the obscuring of a real effect for the wrong reason. Bias results from comparing subjects that differ in some important way. It is the failure to isolate, for a specific risk factor, an accurate measure of effect (separate from random error), and compromises the internal validity of a study. While many types of bias can be defined [Sackett, 1979], typically researchers focus on three specific types of bias: selection bias, information bias, and confounding [Rothman et al., 1998]. For the purposes of this manuscript, we focus on the first two types of bias and evaluate their importance with respect to studies of childhood leukemia and residential magnetic field exposures.

Selection Bias

In designing an epidemiologic study, it is important to assess the characteristics of the two groups to be compared and the possible impact of bias on study results. These characteristics depend on the design of the study. For case–control studies, all subjects (i.e., cases and controls) must be comparable and representative of the general population from which they are drawn. The investigator should select subjects based on each subject's disease status without knowledge of their exposure status. If there is a differential preference between cases and controls for selecting individuals with (or without) exposure, this may induce a bias in the measure of association between disease status and exposure status. For cohort studies, all subjects are disease free at the outset and are selected based on their exposure status without knowledge of their disease history. Provided that the investigator has no knowledge of the subjects' disease history (for historical cohort studies) and that follow-up is complete, there is no selection bias. However, it is essential that case ascertainment be complete and independent of exposure status. If the follow-up of individuals differs by disease status, follow-up bias may occur which can give erroneous results. In cohort studies, the representativeness of the subjects is relevant only for generalizing the study results to other populations (i.e., external validity).

Below are described specific aspects of studies of residential magnetic field exposure and childhood cancer studies that may have led to bias (Tables 1 and 2). This discussion identifies possible sources of bias.

Contract grant sponsor: Public Health Institute, California Department of NIEHS Center; Contract grant number: ES05022.

*Correspondence to: Daniel Wartenberg, 170 Frelinghuysen Road, EOHSI, Piscataway, NJ 08854. E-mail: dew@eohsi.rutgers.edu

Received for review 2 February 2000; Final revision received 9 August 2000

While such bias may exist, it is not possible to determine conclusively from published data whether the biases are large enough to explain the reported associations or lack of associations. Additional data and analyses would be required to determine this. Since all but one of the studies are case–control studies, I do not address issues of cohort studies in any detail. For case control studies, I focus on four main issues: representativeness of sources of cases, representativeness of sources of controls, participation rates, and differential mobility of subjects. Finally, I discuss a new and innovative approach that has been developed to assess possible control selection bias: the case specular method.

Representativeness of Cases

Three epidemiologic designs have been used in the studies of childhood cancer and residential exposure to magnetic field residential exposure studies (Table 1). The first type of study used a standard case–control design. That is, subjects with disease and subjects without disease were identified, and their exposures were estimated retrospectively. Cases were selected from population-based or hospital-based registries. Controls were selected from another register (e.g., birth certificates), through a random digit dialing procedure, or from friends of the cases.

The second type of study was a nested case–control design. In this design, a cohort of individuals was identified containing both cases and controls and both exposed and unexposed subjects, such as those living near electric power transmission lines. From this cohort, cases and controls were chosen independently of their exposure and compared with respect to their exposure as in a traditional case–control study. The advantage of this design is that, if designed properly, it enables the investigator to select a preferable exposure prevalence among subjects compared to the general population (in the case of residential magnetic fields, higher exposure prevalence), increasing the study's sensitivity for detecting an association between exposure and disease.

The third type of study was an historical cohort study. In this study, investigators identified all persons living near the electric transmission facilities as their cohort, and compared the incidence experience of these people with national incidence rates.

All the case–control studies utilized registry records to identify cases. They used all cases diagnosed or dying during a certain range of years, living in a specified geographic region and below a certain age (Table 1). Three different types of data sources were used to identify cases. Wertheimer and Leeper [Wertheimer et al., 1979] used death certificates. One concern with this source is that severity of disease may have affected reporting (i.e., non-fatal leukemias do not result in death certificates). Diagnosis, treatment and access to care can affect disease severity and may be related to socioeconomic status. And, socioeconomic status, in turn, could be related to proximity to power lines, which could have resulted in a bias. Fulton [Fulton et al., 1980], Fajardo-Gutierrez [Fajardo-Gutierrez et al., 1993] and Petridou [Petridou et al., 1997] used hospital incidence registries or physician registries. Socioeconomic status may influence physician and hospital choice. Referral patterns also can affect hospital choice. Unless the controls are drawn from the same population as the cases, i.e., the same hospital or physician, bias may result. The rest of the case–control studies used population-based registries to identify cases [Tomenius, 1986; Savitz et al., 1988; Coleman et al., 1989; Myers et al., 1990; London et al., 1991; Feychting et al., 1993; Olsen et al., 1993; Verkasalo et al., 1993; Linet et al., 1997; Michaelis et al., 1997; Tynes et al., 1997; Dockerty et al., 1998; Green et al., 1999a,b; McBride et al., 1999] These are optimal provided that ascertainment of cases is sufficiently high. Four studies supplemented these registries with other data [Savitz et al., 1988; Myers et al., 1990; Dockerty et al., 1998; McBride et al., 1999]. Overall, while there are minor variations in case selection, published data are not sufficient to assess possible bias. I believe these variations are unlikely to produce a largely unrepresentative sample or a bias large enough to explain the observed results.

If exposure to magnetic fields had an age-specific effect on the incidence of leukemia, studies using children of ages outside those affected could bias the reported effects toward showing no association. The age ranges of cases in studies reviewed here varied, from allowing only subjects under 11 years of age to allowing all those under 21. These ranges all are representative of children. Most studies used children under 15 as those eligible. Age variation is unlikely to have caused a substantial bias since there was such a large overlap between the age ranges, although it would be worthwhile to investigate the sensitivity of the results to maximum age cutpoint. Similarly, the calendar year ranges of eligibility vary across studies but should not result in substantial bias. Again, a sensitivity analysis could produce some insight. These sensitivity analyses cannot be conducted thoroughly without the original data.

Representativeness of Controls

In case–control studies, controls should provide an estimate of the exposure distribution in the population from which the cases were drawn [Rothman et al.,

TABLE 1. Subject Selection in Residential Childhood Cancer Studies

		Source of subjects		Eligibility			
Design	Reference	Cases	Controls	Years	Location	Ages	Matching criteria
Case–control	Wertheimer	Death certificates	Next birth certificate unless a sibling	1950–1973	Colorado birth and Greater Denver Resident 1946–1973	<19	Either month and county of birth (File 1) or alphabetical, 5–20 year range (not sibling) (File 2)
	Fulton	State hospital incidence registry	Birth certificate	1964–1978	Rhode Island resident for 8 years prior to diagnosis	<21	Birth year
	Tomenius	Population-based cancer registry	Nearest birth certificate in Parish records	1958–1973	Born and diagnosed in Stockholm County	<19	Age, gender, church district of birth, and church district of diagnosis (if same as birth for case)
	Savitz	Population-based cancer registry and hospital records	Random digit dialing	1976–1983	Denver SMSA	<15	Age ±3 years, gender, telephone exchange at time of diagnosis of case
	Coleman	Population-based cancer registry (leukemias only)	Solid tumors (not lymphomas)	1965–1980	4 London Boroughs	All ages	Age, gender, year of diagnosis, residence borough
	Myers	Population-based cancer registry and other sources (mortality)	Nearest birth certificate	1970–1979	Yorkshire Health Region	<15	Gender, birth in same local area or health district, year of diagnosis
	London	Population-based cancer registry (leukemia)	Friends (first 65) and random digit dialing	1980–1987	Los Angeles County	<11	For most, age ± 1–3 years depending on age, gender, ethnicity
	Olsen	Population-based cancer registry (leukemia, CNS tumor, lymphoma)	Population registry	1968–1986	Denmark	<15	Gender, date of birth ± 1 year
	Fajardo-Gutierrez	Hospitals (3rd level)	Hospitals (3rd level)	—	Mexico City	—	Same hospital (in patient or out patient); no neoplasm
	Linet	Population-based cancer registry	Random digit dialing	1989–1994	9 U.S. States (IlL, IN, IO, MI, MN, NJ, OH, PA, WI)	<15	Phone number (8 digits), age, race
	Michaelis	Population-based cancer registry	Population registry	1991–1994	Berlin, Germany; Lower Saxony	<15	Gender, data of birth, city district at diagnosis
	Petridou	Physicians network	Hospital records	1993–1994	Greece (Greek nationals)	<15	Gender, age, town (urban) or region (rural), hospitalized at same time

(*Continued*)

TABLE 1. (*Continued*)

	Dockerty	National cancer registry hospital admission/ discharge data children's cancer registry	National birth records	1990–1993	Born and diagnosed in New Zealand	<15	Gender, age, resident in New Zealand
	McBride	Pediatric oncology treatment centers population-based cancer registry	Provincially-based government health insurance rolls	1990–1994	Within 100 km of principal cities of British Columbia, Alberta, Saskatchewan, Manitoba, Quebec	<15	Gender, age, area
	Green	Hospital for sick children via the pediatric oncology group registry	Randomly selected from telephone marketing lists	1985–1993	Resident within metropolitan Toronto or counties of York, Durham or Peel	<15	Gender, year of birth
Nested case–control	Feychting	Population-based cancer registry	Population registry	1960–1985	Residence within 300 m of any 220 kV or 400 kV power line in Sweden	<16	For most, in registry during year of diagnosis, birth year, gender, residence in same parish in year of diagnosis or move, near same power line
	Tynes	Population-based cancer registry	Population registry	Selected years 1960–1989	Residence in a census ward crossed by a high voltage power line during at least one of: 1960, 1970, 1980, 1985, 1987, 1989	<15	Vital status at time of diagnosis, sex, year of birth, municipality (except a few selected from neighboring municipality)
Historical cohort	Verkasalo	National cancer registry	National population registry	1970–1989	With 500 m overhead power lines	<20	5-year age groups

TABLE 2. Numbers of Subjects in Residential Childhood Cancer Studies

Reference	Number of subjects originally identified		Excluded/refused		Moved		
	Leukemia cases	Controls	Cases	Controls	Cases	Controls	Comments
Wertheimer and Leeper	344 (155 leukemia?)	344 (155 controls)	16—no death address; 72—no birth address	16—no death address; 72—no birth address	147	128	21% HCC in File 1; 23% HCC in File 2
Fulton et al.	119	240	9	15	53	—	Cases had to live in RI for 8 years prior to diagnosis; Controls had to be born in RI
Tomenius	746 (56 benign)	716	29—bad data; 1—not primary tumor; 43 of 1172 dwellings	46 of 1015 dwelling	316 of 746		
Savitz et al.	103	278	30 interviews; 67 measurements; 6 wire codes	56 interviews; 71 measurements; 19 wire codes (78.9% response rate in RDD)	37?	0	Controls had to be resident in Denver at time of recruitment and time of diagnosis of case; mobility of cases greater than that of controls
Coleman et al.	811 (84 under age 18)	1614 (141 under age 18) 254	40	182 23	—	—	Resident at time of selection (or 1975 for 2nd controls) used; mobility not assessed
Myers et al.	419	656	45	68	NR	NR	Analyses based on birth residence; control residence at time of diagnosis not reported
London et al.	331	257	99 interviews; 162 measurements; 112 wire codes	24 interviews; 108 measurements; 50 wire codes (82% response rate for RDD)	57%	66%	4,424 phone numbers resulted in 113 eligible controls
Olsen et al.	1707	4788	0	0	1050	3125	
Fajardo-Guiterrez	81	77	—	—	—	—	Little ancillary information reported
Linet	942	1,292	304 no interviews	672 no interviews	66%	68%	Used complete residential histories; mobility similar between cases and controls

(*Continued*)

TABLE 2. (*Continued*)

Michaelis	283	919	43 questionnaire; 64 measurement	339 questionnaire; 166 measurement	—	—	No historical residency requirement
Petridou	117	202	0	14 (8 replaced)	—	—	Require to be at same residence for at least 4 years
Feychting and Ahlbom	142	558	1 calculations; 53 measurements	4 calculations; 214 measurements	—	—	Residence history used in exposure assessment
Tynes	532 (148 leukemias)	2112 (579 controls)	32	108	NR	NR	Had complete residential history; excluded those not living in wards prior to diagnosis
Verkasalo	32	—	—	—	—	—	Cohort study
Dockerty	344 (131 leukemias)	—	121 interviewed; 115 measured	121 included (secondary controls); 117 measured	75	77	40 matched pairs resident for at least 2 years prior to diagnosis date analyzed separately
McBride	449	675	4 not located: 46 not contacted: about 75% measured	149 not located; 127 not sued; about 85% measured	NR	NR	Mobility of cases was higher than that of controls
Green	256	645	22 physician refusal; 22 relocated outside study area; 36 no response; 17 refused; 2 discarded	45 no response; 168 refused; 13 discarded	NR	NR	Mobility of cases was higher than that of controls

1998]. They should provide an estimate of the exposure rate that would have been observed in the cases if there were no association between the exposure under study and the disease [Schlesselman, 1982]. Controls (or comparison populations) were selected in a variety of ways. Some studies used regional birth certificate files [Wertheimer et al., 1979; Fulton et al., 1980; Tomenius, 1986; Myers et al., 1990; Dockerty et al., 1998]. This limits subjects to those who were both born and diagnosed (or selected) in the same region. In general, this is advisable to increase the likelihood adequate duration of residence (i.e., exposure) at the specified house.

Fulton [Fulton et al., 1980] conducted a hospital-based case–control study but used the general population listed in the birth certificate records as the source for controls. This population likely was larger than the hospital population from which the cases were drawn, which could have led to bias. Further, Wertheimer and Leeper [Wertheimer et al., 1980] have argued that because matched control homes were selected by birth addresses while case homes (often more than one per case) were homes occupied any time 8 years prior to diagnosis, there is a deficit of suburban addresses in the control population (or an excess of urban addresses). This resulted in a bias towards higher exposure for controls. In addition, only address at birth was used for controls while complete case address histories were obtained and used.

Other studies used random digit dialing to identify controls [Savitz et al., 1988; London et al., 1991; Linet et al., 1997]. Random digit dialing is a method designed to identify a set of controls for a study that come from a defined geographic region [Waksberg, 1978; Robison et al., 1984; Ward et al., 1984; Greenberg, 1990; Olson et al., 1992; Voigt et al., 1992; Lele et al., 1994; Psaty et al., 1994; Brick et al., 1995; Sakkinen et al., 1995]. For each case identified, the investigator takes the case's phone number, discards the last two digits and replaces them with two randomly chosen digits. This number is called. If it is not a residence, another pair of random digits is used. If it is a residence, the interviewer asks if a subject meeting the matching criteria resides there. If so, this person is recruited as a subject. If not, another pair of random digits is used. This process, applied to a childhood study with gender, age and ethnic matching, typically requires between 25 and 75 phone calls per case to identify an eligible control.

One limitation of random digit dialing is that it samples only homes with telephones. It is important to determine what proportion of residences have telephones and, if possible, to compare those residences with those that do not with respect to exposure and confounding variables. Poole and Trichopoulos [Poole et al., 1991] argue further that people of very low socioeconomic status are harder to reach by this method and are underrepresented in the sample. Another limitation of random digit dialing is that there is limited ability to assess non-response. Investigators rarely have information about residents at telephone numbers that never respond. More recently, as the use of telephone lines for fax machines, computer lines and cellular phones has increased, the effort required to reach the owner of the phone line has increased dramatically and the representativeness of those reached as a random sample has become increasingly questionable.

Gurney and colleagues conducted a study in the Seattle area to evaluate the potential bias of random digit dialing with respect to a possible association between socioeconomic status and wire codes [Gurney et al., 1995]. That study found that high wire code homes were more likely to be of low income families, although the association was weak. Since high income is associated with increased risk of childhood leukemia, the finding is consistent with a downward bias of the true odds ratio. If, however, the control homes selected were of higher socioeconomic status, due to more variation between than within dialing regions, an upward bias would be seen. Documenting this bias is complicated although the effect is unlikely to be large.

Another possible source of bias in using random digit dialing is that the sampling unit is the residence rather than the individual. That is, the investigator tries to reach each residence via telephone rather than each individual child, as one does in a registry sampling procedure. Individuals are the unit used in the final analysis. But if a residence does not have a telephone, none of the children in that residence can be considered for use as a control or, if a home has more than one eligible child, typically only one is considered eligible for the study. Exclusion of homes without phones likely results in greater similarity among potential subjects, while limiting subjects to one per household decreases similarity among potential subjects.

One study selected some controls using random digit dialing and others who were friends of cases [London et al., 1991]. The latter approach involves asking a case to name a friend who could be approached for inclusion in their study. One of the problems in using friend controls is that they may be overmatched [Kelsey et al., 1996]. That is, friend controls may be more similar to the cases in terms of a factor that is not a confounding variable, but is associated with exposure, compared to population-based controls. This reduces the statistical efficiency, meaning that a

larger sample size might be needed to detect the association of interest, because the exposures tend to be more similar than expected (under a random control selection scheme) due to the friend-matching [Rothman et al., 1998]. It also can induce confounding which, if not adjusted for, can create bias. In addition, since the name of the friend is solicited from the case (or surrogate), it is possible that the case (or surrogate) has exercised some type of selection bias, such as selecting the friend that is most talkative or out-going, which may in turn be related to some risk factors. If one wishes to use friends as controls, one could solicit a list of friends from each case and select the controls for a case at random from such a list. Even so, there may be some bias. In addition, selecting some controls by one method and others by another method may lead to heterogeneity among controls. The rationale of combining the two different methods had to do with logistical considerations. It still may have led to differences between controls (i.e., different chances for selection depending on the method). Analyses of the data stratified or separated by control selection method could help investigators evaluate this.

Another study selected controls randomly from a marketing list and then contacted them by phone [Green et al., 1999a, b]. This is even more problematic than random digit dialing as it starts from a selective and likely biased list (at least in terms of socioeconomic status) and then encounters many of the same access limitations as random digit dialing.

One study used two complete sets of controls. One was composed of cancer cases other than leukemia and lymphoma, and the other set of controls was drawn from the local electoral roll [Coleman et al., 1989]. Both of these control populations are problematic. Other cancer cases may lead to a negative bias (reduction in the possible association) if cancers other than those excluded (i.e., brain cancers) are associated with exposure to magnetic fields. Fortunately, in this case, there are scant data to support an association of childhood cancers other than leukemia or brain cancer with magnetic field exposures. The electoral roll, which was not used for the childhood portion of the study, may not include all persons living in the region and may be ethnically and socioeconomically biased.

Still another study selected controls for a government health insurance listing [McBride et al., 1999]. Issues regarding completeness would have to be considered.

The cohort study, two case–control studies, and both nested case–control studies used population registries to identify controls [Feychting et al., 1993; Olsen et al., 1993; Verkasalo et al., 1993; Michaelis et al., 1997; Tynes et al., 1997]. Again, if ascertainment is adequate, this is ideal as samples are drawn at random from the entire population that one wishes to sample. Often, the investigator does not have access to such a convenient database.

Finally, two studies used the residence as the unit of analysis [Fulton et al., 1980; Tomenius, 1986]. This may introduce bias because the number of residences per subject may vary by disease and/or exposure (i.e., residence at birth, residence at diagnosis or death). Further, if the total number of residences is used in any of the analyses, the apparent sample size is increased artificially (i.e., more than one home per subject) inappropriately increasing the precision.

Participation Rates

Having identified the source of subjects, the next major concern in the case–control studies is the possible bias due to non-participation. If non-participation rates differ by exposure status only or by disease status only, there is no bias in the odds ratio. However, if these rates differ by both exposure and disease, a substantial bias may exist. One can conduct a sensitivity analysis for such an effect to determine the maximum amount of bias that could be present.

The potential for this problem is shown in Table 2. Some of the studies show substantial numbers of exclusions or refusals to participate. Without characterizing these individuals, it is not possible to determine the degree of bias imparted. However, if one were to do a sensitivity analysis to determine how large an effect there might be for these exclusions, in worst case, it is likely that the exclusion adjusted odds ratios would be substantially different from those reported. While this problem is not unusual in epidemiologic studies, it can have a large impact.

Gurney and colleagues, in their study of the potential bias from the use of random digit dialing for control selection, also evaluated the possible role of participation bias. They assumed that low income controls were less likely to participate in the study than other controls. Using their data on wire code distributions and income and assuming that all cases participated, they estimated that if 80% of controls with incomes over $15000 participated in a study and 0% with incomes under $15000 participated, it would produce a bias in the odds ratio of 1.24. If participation were better in those of low income or worse in those of high income, the effect would be even smaller. Again, these data suggest a possible bias, but not one large enough to explain the observed results.

The problem of participation bias is of particular concern across the range of measurement methods used. In most studies, far more houses were assessed in terms of wire codes than magnetic field measures. It

would be instructive to get the data on the homes in all of these studies and determine if there was a bias associated with the type of exposure measurements made. For example, Savitz [Savitz et al., 1988] reports that homes with missing measurement data were more likely to have high or very high current configuration wire codes than those with measurements (28.8% vs. 22.1%), and there were more homes missing measurements among cases than controls. Thus, for that one study, the missing data likely resulted in an underestimation of the size of the association. Similarly, Hatch [Hatch et al., 2000] reports that subjects who had front door magnetic field measurements but not indoor magnetic field measurements, had higher magnetic fields (15.6% vs. 12.7%) and more of the highest category of wire codes (8.8% vs. 6.3%) than those with both measurements. Based on limited data from these two studies, subjects in homes with fewer types of measurements had higher exposures.

Mobility

Differential mobility of cases and controls has been raised as another possible source of bias in these studies. While generally data needed to evaluate this possibility are not published, some of the studies do present information on how many residences have been occupied by each study subject (Table 2). Jones and colleagues argue that the observed associations between wire codes and childhood cancers may be due to bias induced by differential mobility (for example, in the Savitz study, controls were required to be residentially stable but cases were not), and mobility has been found to be associated with higher wire codes [Jones et al., 1993]. Mobility differences were evident in the studies of Tomenius [Tomenius, 1986], Savitz [Savitz et al., 1988], McBride [McBride et al., 1999], and Green [Green et al., 1999a, b].

In most studies where it was reported, cases were more mobile than controls. In part, this may be due to different stability restrictions on the eligibility of cases versus controls. The purpose of a residential stability restriction in one study design was to eliminate from potential controls those who were not resident at the time of case diagnosis and who thus would not have been a case even if they had developed disease [Savitz et al., 1988]. In a study to investigate the possibility of this phenomenon, Jones found 31% more high wire codes in non-stable than in stable populations [Jones et al., 1993]. Again, while plausible, the quantitative impact is limited. Only if one assumes that the 31% excess of high wire code found by Jones [Jones et al., 1993] were true for virtually all of the Savitz study cases and none of the controls [Savitz et al., 1988], would removing this excess lower the odds ratio towards showing no association. Further, the applicability of the Jones data to the Savitz data is questionable since the Jones et al. study had substantially more high wire codes overall (30.1% vs. 24.5% total; 34% vs. 28% cases; 26% vs. 20% controls) [Jones et al., 1993].

Reviews

Issues of selection bias have been raised in various reviews of these residential childhood cancer studies. For example, the NRPB Report [National Radiological Protection Board, 1992] raised issues of bias in the use of random digit dialing for both the Savitz [Savitz et al., 1988] and London [London et al., 1991] studies. They suggested that this results in an undersampling of controls with low income and in differential mobility between cases and controls for these two studies.

The ORAU Report [Oak Ridge Associated Universities, 1992] reviewed control selection bias first for studies they viewed as most important [Wertheimer and Leeper, 1979; Savitz et al., 1988; London et al., 1991] and then for the others [Fulton et al., 1980; Tomenius, 1986; Coleman et al., 1989; Myers et al., 1990]. The authors argued that the control selection procedure used by Wertheimer and Leeper [1979] was not defined with sufficient clarity for critical evaluation, although they acknowledged it contained no obvious bias. They also noted that, for the Savitz et al. study [Savitz et al., 1988], if exposure were related to residential mobility or the chance of being sampled as a control, control selection bias would have been introduced. The underrepresentation of those of lower socioeconomic status in random digit dialing was raised with reference to the Savitz and London studies [Savitz et al., 1988; London et al., 1991]. They also questioned the representativeness of using friends as controls in the London et al. [1991] study.

For those studies deemed less important, the ORAU report also identified issues of control selection bias. For Fulton, they pointed out that cases had to have residential stability (residence near a specific hospital) while controls did not (residence anywhere in Rhode Island) [Fulton et al., 1980]. Coleman et al. used cancer controls in their study, which may have biased the result downward since brain cancer cases, which are the second most common childhood cancers, were acceptable controls and may have been associated with exposure to magnetic fields [Coleman et al., 1989].

The NRC and NIEHS Working Group reports raised many of the same issues as stated here [National Research Council, 1997; NIEHS Working Group, 1998]. Their conclusions were that, ''empirical efforts to characterize the potential bias yielded plausible and

testable hypotheses regarding social class, nonresponse and residential stability, but little direct support that the bias actually occurred.'' In addition, the NIEHS report noted that if higher socioeconomic status was associated with leukemia and random digit dialing was used to select controls, the study was likely to have overestimated risk. Similarly, if controls were more residentially stable than cases, the study may have overestimated risk.

Case Specular Method for Assessing Control Selection Bias

Control selection bias has been cited by many as a possible explanation for the observed association between wire codes, magnetic fields and childhood cancer. While many aspects of this issue have been discussed, as noted above, the data available preclude a convincing assessment. To examine this issue from another perspective, Zaffanella and colleagues developed an innovative approach using existing data and new wire code evaluations of properties near those of the study subjects [Zaffanella et al., 1995, 1998a]. The method was designed to discriminate between two hypotheses: (1) childhood cancer is associated with magnetic fields and wire codes are a surrogate for magnetic fields; and (2) childhood cancer is associated with neighborhood factors other than magnetic fields, such as socioeconomic status or environmental agents such as air pollution, and wire codes are a surrogate for these neighborhood factors. In this method, a control home is defined as a hypothetical (virtual or specular) home directly across the street from each case home. The hypothetical control house is a mirror image of the case house, thus matching on many neighborhood characteristics but not necessarily with the same wire code configuration. The wire codes are assessed for the case and specular homes and a summary measure of effect is calculated. The underlying concept of this method is that if high wire codes are an independent risk factor for cancer, there should be more cases living on streets with higher wire codes, and more cases living on the side of the street where the power lines are located. If, on the other hand, the cancers arise from some other neighborhood risk factor, the wire codes of the case and specular homes should be similar.

There is one major assumption underlying this method, as noted by the authors: that residences on one side of the street are not systematically different from residences on the other side of the street except possibly for wire codes. This assumption is not validated. One commenter [Maclure, 1998] explains that he, ''thinks of a city as a mountainous region in which the peaks, valleys and slopes are the levels and gradients of socioeconomic and unknown factors that influence where houses are located and who lives in them. The streets are like rivers with relatively symmetrical banks.'' Therefore, houses directly across the street from each other, he asserts, are more like each other than houses one or two doors up or down the street. Again, this is an assumption and is not validated.

The case specular method has been applied in conjunction with data from two studies [Zaffanella et al., 1997; Zaffanella et al., 1998b; Ebi et al., 1999]. The results depend on the discordance of the specular pairs, i.e., when the wire code of the residence is different from the wire code of its specular. For the Denver data, the discordance rate was 34% while for the Los Angeles study it was 50%. The results, summarized in Table 3, showed that the odds ratios using the speculars as controls were higher than with the original controls, suggesting that there was not another neighborhood factor, such as socioeconomic status, responsible for the observed associations. Further, the observation that odds ratios increased when specular controls were used instead of random digit dialing controls is consistent with the hypothesis that cases tended to live on streets with higher wire codes.

It is important to note that this analysis did not address this issue of whether there was control selection bias in the data. That bias may still have been present, but this analysis was inconsistent with confounding by a neighborhood factor.

In summary, while various aspects of potential control selection bias have been identified, the series of issues investigated herein have provided little data to support or validate the claim that they were responsible for the observed elevations in odds ratios.

Information Bias

Another type of bias, which I will treat more briefly, is information bias. Information bias occurs when there are errors in the information gathered about study subjects. Errors can be made in two ways: at

TABLE 3. Odds Ratio (95% Confidence Interval) From the Application of the Case Specular Method (HCC vs. LCC)

	Case–control	Case-specular	Case-specular control-specular
Savitz: childhood cancer	1.6 (1.1–2.3)	2.3 (1.3–3.9)	2.4 (0.9–6.1)
London: childhood leukemia	1.5 (1.0–2.2)	1.8 (1.1–3.0)	2.4 (1.2–4.9)

random or preferentially in one group of subjects. If made at random, the errors are termed non-differential, i.e., they do not differ in occurrence between the two groups. These errors dilute any observed effect, giving results that show less of an association between exposure and disease than actually exists. In contrast, errors may occur preferentially between cases and controls or exposed and unexposed subjects. This is called differential misclassification. Bias may cause either weaker or stronger associations to be observed than actually exist, depending on the direction of the misclassification.

For the studies under consideration, there are two main types of information: disease; and exposure. In most cases, disease data have been confirmed histologically and are unlikely to have resulted in such consistent results. Exposure, on the other hand, has been evaluated in several ways and is much more complex.

Exposure Assessment

Our concern in this study is residential exposure assessment. The basic issue with exposure assessment is whether exposure for individual subjects is measured accurately. If not, it is important to know whether the errors are non-differential, i.e., made irrespective of the disease status, or differential. While we cannot assess this directly, we can review the issue in general, noting specific limitations of the methodology used. Note that mention of measurements includes (1) magnetic field measurements, (2) measurements of the distance of the house from the power lines, and (3) calculated fields which are based on distance measurements and measurements of the electrical load on the lines of the power delivery system.

One major limitation of the exposure assessments of the magnetic field and childhood cancer studies is that exposure assessments were limited to the exposures which occurred in the home, with the exception of one study that looked at school exposures [Lin et al., 1994]. Exposures in the home come from several sources: electrical facilities outside the home; wiring to and in the home including current on water pipes, etc.; and electrical appliances, toys, etc. in the home. Based on available evidence, the same source is not dominant in all homes. Use of residential exposure only could result in misclassification. Large facilities outside the house often use large amounts of electrical power and may give rise to greater magnetic fields than typical residential wiring, e.g., feeder lines, elevators, large electrical equipment such as generators and industrial size appliances. Similarly, if children spend substantial amounts of time in daycare, at their parent's workplace, or elsewhere outside the home, exposures may vary greatly.

The second major limitation of the exposure assessments has been to limit estimates to magnetic fields from the electrical power distribution system. Most studies estimated the subject's exposure from external sources by examining proximity of the home to the electrical facility. Because observable markers like number of conductors and gross estimates of their size are very crude indicators of line current and because some important highly-variable parameters, such as net and ground currents have no observable marker, wire codes can only be viewed as a very crude indicator of magnetic field within a nearby home, as demonstrated by the wide dispersion in field levels measured in homes with similar wiring codes [Zaffanella, 1993]. Exposure to appliances and sources outside the house were rarely considered.

A third limitation of the exposure assessments is the lack of consistent data on the relative size and characteristics of estimation errors. Since we lack consistent data on those errors, there is little scientific basis on which to argue the relative merits of measurements and wire codes. One might intuitively expect measurements which quantify fields from all sources, to be a better indicator of overall residential exposure than methods which use crude estimates from only one source. But contemporary short-term measurements offer two more sources of error: the error between the short-term measurement and contemporary long-term average, and the error between contemporary average and the average field at the historical exposure time of interest. We also have no information about the relative size of errors in magnetic field measurements in comparison to distance measurements or line load measurements.

The fourth limitation of the exposure assessments is the other biases in the measurements that are not always apparent. Principally, these biases have involved measurements in the center of the room and measurements in ''low power'' conditions. Both of these tend to bias exposure estimates toward measurements of fields from exterior sources by avoiding more localized fields from wiring in the floors, walls, and ceilings or fields from appliances that tend to be arranged around the periphery of the room. Not surprisingly, therefore, ''low power'' center of room measurements correlate with wire codes better than measurements in more realistic locations under more realistic conditions. Similarly, distance measurements, which are used to estimate exposure directly and also in conjunction line load measurements, are based on distance of the line from the house and do not differentiate among the different locations within the dwelling.

The fifth limitation of exposure assessments is that electric and magnetic fields are ubiquitous.

Everyone is exposed to some degree, making it extremely difficult to define a reference or ''no exposure'' population. This likely results in a downward bias of any measured effect.

Finally, we must note that, in addition to how exposure is measured, the way in which investigators summarize exposure data can lead to bias. That is, when categorizing continuous magnetic field data investigators often make arbitrary decisions about category boundaries, and these boundaries may differ from study to study. This results in cutpoint bias, the use of a non-representative odds ratio due to the choice of exposure classification rule [Wartenberg et al., 1991a, b]. Elsewhere, we have shown that this occurs with magnetic field data and can lead to different interpretations of results [Wartenberg et al., 1993]. One approach to this problem is to use continuous exposure-response models to summarize the relationship [Greenland, 1995]. Short of that, one can strive to use the same or similar cutpoints for consistency. In that way, results of studies using similar cutpoints can be compared directly. But, any categorization of continuous decreases precision and can lead to bias.

In summary, it is important to remember there is no persuasive evidence that one approach to exposure assessment is necessarily better than another. There are some studies in which exposure assessment is carried out poorly and they can be identified. But, there is no basis to maintain that measurements are better than wire codes or vise versa. When reporting results, we should use continuous exposure-response models, when possible, and consistent exposure cutpoints otherwise. Finally, we have no evidence to suggest that errors in exposure assessment are differential. If not, errors will weaken any true association between exposure and disease.

CONCLUSION

This paper reviews the main sources of bias in studies of residential exposure to magnetic fields and childhood cancer. The issues addressed include study design, selection of cases, selection of controls, participation rates, mobility, and exposure assessment. All of these can affect the results of a study. Unfortunately, published data rarely allow readers to evaluate the role of most of these. What is most compelling is when a body of literature includes a wide diversity of designs and methods, so that one can look for consistency to evaluate possible bias. Unfortunately, this is not always the case. For example, this paper reports on 18 studies of childhood leukemia and exposure to magnetic fields, of which only one was not a case–control study. Thus, all the biases associated with case–control studies may be present. The lack of cohort studies is likely due to the rarity of both the disease and the exposure in the general population, making conduct of such a study extremely expensive and statistically under powered. It is fortunate that the studies were conducted in different regions of countries, in different countries, and in different years. Concordance of results from studies showing wide variation in the populations studied suggests only limited bias.

After study design, control selection bias is probably the next biggest concern in these studies. We have limited information on source of controls, mobility of subjects, and participation rates of subjects. While these data do not reveal any striking patterns that suggest strong bias, far more information is needed to be convincing. In addition, a newly developed method to assess possible control selection bias provides data that suggest that if a bias exists it is a bias toward the null.

Concerns over information bias are also important to evaluate. Information bias can occur through errors in disease reporting and exposure measurement. There are few publicly available data to assess this directly. Again, data exist to suggest some small effects, but nothing shows obviously large effects. These may result in some bias, and certainly result in a reduction in precision.

As with any epidemiologic study, the studies of residential magnetic field exposure and childhood cancer have many possible sources of bias. Unfortunately, these are very hard to quantify unless we systematically analyze the data, which generally are not available. While any of these biases could contribute to the size of the reported odds ratios, none is believed to be so substantial that in all of these studies it could be used to explain the results. Rather, each may contribute in a small way to the odds ratio, either increasing or decreasing its value. Since there is such variety in the populations studied and because there is no overriding flaw that encompasses all of them, it seems unlikely that selection bias cannot be the sole explanation for the reported associations between exposure to magnetic fields and childhood cancer incidence.

ACKNOWLEDGMENTS

This project was funded by Electric and Magnetic Fields Program of the Public Health Institute and the California Department of Health Services. I thank many colleagues for sharing the views on these issues with me, particularly Fred Dietrich, Kris Ebi, Robert Goldberg, Charles Poole, David Savitz and Nancy Wertheimer.

REFERENCES

Brick JM, Waksberg J, Kulp D, Starer A. 1995. Bias in list-assisted telephone samples. Public Opinion Quarterly 59:218–235.

Coleman M, Bell C, Taylor H, Primic-Zakelj M. 1989. Leukaemia and residence near electricity transmission equipment: A case–control study. Br J Cancer 60:793–798.

Dockerty JD, Elwood JM, Skegg DCG, Herbison GP. 1998. Electromagnetic field exposures and childhood cancers in New Zealand. Cancer Causes Control 9:299–309.

Ebi KL, Zaffanella LE, Greenland S. 1999. Application of the case-specular method to two studies of wire codes and childhood cancers. Epidemiology 10:398–404.

Fajardo-Gutierrez A, Garduno-Espinosa J, Yamamoto-Kimura L, Hernandez-Hernandez D, Gomez-Delgado A. 1993. Close residence to high electric voltage lines and its association with children with leukemia. Bol Med Hosp Infant Mexico 50:32–38.

Feychting M, Ahlbom A. 1993. Magnetic fields and cancer in children residing near Swedish high-voltage power lines. Am J Epidemiol 138:467–481.

Fulton J, Cobb S, Preble L, Leone L, Forman E. 1980. Electrical wiring configurations and childhood leukemia in Rhode Island. Am J Epidemiol 111:292–296.

Green LM, Miller AB, Agnew DA, Greenberg ML, Li J, Villeneuve PJ, Tibshirani R. 1999. Childhood leukemia and personal monitoring of residential exposures to electric and magnetic fields in Ontario, Canada. Cancer Causes Control 10:233–243.

Green LM, Miller AB, Villeneuve PJ, Agnew DA, Greenberg ML, Li J, Donnelly KE. 1999. A case–control study of childhood leukemia in southern Ontario, Canada, and exposure to magnetic fields in residences. Int J Cancer 82:161–170.

Greenberg ER. 1990. Random digit dialing for control selection: A review and a caution on its use in studies of childhood cancer. Am J Epidemiol 131:1–5.

Greenland S. 1995. Dose-response and trend analysis in epidemiology: Alternatives to categorical analysis. Epidemiology 6:356–365.

Gurney J, Davis S, Mueller B, Kaune W, Stevens R. 1995. Childhood cancer occurrence in relation to power line configurations: A study of potential selection bias in case–control studies. Epidemiology 6:31–35.

Hatch EE, Kleinerman RA, Linet MS, Tarone RE, Kaune WT, Auvinen A, Baris D, Robison LL, Wacholder S. 2000. Do confounding or selection factors of residential wiring codes and magnetic fields distort findings of electromagnetic fields studies? Epidemiology 11:189–198.

Jones TL, Shih C, Thurston D, Ware B, Cole P. 1993. Selection bias from differential residential mobility as an explanation for associations of with childhood cancer. J Clin Epidemiol 46:545–548.

Kelsey J, Whittemore AS, Evans AS, Thompson WD. 1996. Methods in Observational Epidemiology. New York: Oxford University Press, 432 p.

Lele C, Holly EA, Roseman DS, Thomas DB. 1994. Comparison of control subjects recruited by random digit dialing and area survey. Am J Epidemiol 140:643–648.

Lin RS, Lee WC. 1994. Risk of childhood leukemia in areas passed by high power lines. Rev Environ Health 10:97–103.

Linet MS, Hatch EE, Kleinerman RA, Robison LL, Kaune WT, Friedman DR, Severson RK, Haines CM, Hartsock CT, Niwa S, Wacholder S, Tarone RE. 1997. Residential exposure to magnetic fields and acute lymphoblastic leukemia in children. N Engl J Med 337:1–7.

London S, Thomas D, Bowman JD, Sobel E, Cheng T-C, Peters J. 1991. Exposure to residential electric and magnetic fields and risk of childhood leukemia. Am J Epidemiol 1349:923–937.

Maclure M. 1998. Reply (Letter). Epidemiology 9:476–477.

McBride ML, Gallagher RP, Theriault G, Armstrong BG, Tamaro S, Spinelli JJ, Deadman JE, Fincham S, Robson D, Choi W. 1999. Power-frequency electric and magnetic fields and risk of childhood leukemia in Canada. Am J Epidemiol 149:831–842.

Michaelis J, Schuz J, Meinert R, Zemann E, Grigat J-P, Kaatsch P, Kaletsch U, Miesner A, Brinkmann K, Kalkner W, Karner H. 1997. Combined risk estimates for two German population-based case–control studies on residential magnetic fields and childhood acute leukemia. Epidemiology 9:92–94.

Myers A, Clayden A, Cartwright R, Cartwright S. 1990. Childhood cancer and overhead powerlines: A case–control study. Br J Cancer 62:1008–1014.

National Radiological Protection Board. 1992. Electromagnetic Fields and the Risk of Cancer. Chilton, Didcot, UK: National Radiological Protection Board. vol. 3, 138 p.

National Research Council. 1997. Possible Health Effects of Exposure to Residential Electric and Magnetic Fields. Washington, DC: National Academy Press, 356 p.

NIEHS Working Group. 1998. Assessment of health effects from exposure to power-line frequency electric and magnetic fields. Research Triangle Park, NC: National Institute of Environmental Health Sciences, National Institutes of Health, U.S. Department of Health and Human Services, Public Health Service. NIH Publication No. 98-3981, 508 p.

Oak Ridge Associated Universities. 1992. Health Effects of Low-Frequency Magnetic Fields. Oak Ridge, TN: The Committee on Interagency Radiation Research and Policy Coordination. ORAU 92/F8.

Olsen JH, Nielsen A, Schulgen G. 1993. Residence near high-voltage facilities and the risk of cancer in children. Br Med J 307:891–895.

Olson SH, Kelsey JL, Pearson TA, Levin B. 1992. Evaluation of random digit dialing as a method of control selection in case–control studies. Am J Epidemiol 135:210–222.

Petridou E, Trichopoulos D, Kravaritis A, Pourtsidis A, Dessypris N, Skalkidis Y, Kogevinas M, Kalmanti M, Koliouskas D, Kosmidis H, Pangioutou JP, Piperopoulou F, Tzortzatou F, Kalapothaki V. 1997. Electrical power lines and childhood leukemia: A study from Greece. Int J Cancer 73:345–348.

Poole C, Trichopoulos D. 1991. Extremely low-frequency electric and magnetic fields and cancer. Cancer Causes Control 2:267–276.

Psaty BM, Cheadle A, Koepsell TD, Diehr P, Wickizer T, Curry S, VonKorff M, Perrin EB, Pearson DC, Wagner EH. 1994. Race- and ethnicity-specific characteristics of participants lost to follow-up in a telephone cohort. Am J Epidemiol 140:161–171.

Robison LL, Daigle A. 1984. Control selection using random digit dialing for cases of childhood cancer. Am J Epidemiol 120:154–156.

Rothman KJ, Greenland S. 1998. Modern Epidemiology. Philadelphia: Lippincott-Raven Publishers, 737 p.

Sackett D. 1979. Bias in analytic research. J Chronic Dis 32: 51–63.

Sakkinen PA, Severson RK, Ross JA, Robison LL. 1995. Random-digit dialing for control selection in childhood cancer studies: The geographic proximity and demographics within matched sets. Am J Public Health 85:555–557.
Savitz DA, Wachtel H, Barnes FA, John EM, Tvrdik JG. 1988. Case–control study of childhood cancer and exposure to 60-Hz magnetic fields. Am J Epidemiol 128:21–38.
Schlesselman J. 1982. Case–Control Studies: Design, Conduct, Analysis. New York: Oxford University Press.
Tomenius L. 1986. 50-Hz electromagnetic environment and the incidence of childhood tumors in Stockholm County. Bioelectromagnetics 7:191–207.
Tynes T, Haldorsen T. 1997. Electromagnetic fields and cancer in children residing near Norwegian high-voltage power lines. Am J Epidemiol 145:219–226.
Verkasalo P, Pukkala E, Hongisto M, Valjus J, Jarvinen P, Heikkila K, Koskenvuo M. 1993. Risk of cancer in Finnish children living close to power lines. Br Med J 307:895–899.
Voigt LF, Davis S, Heuser S. 1992. Random digit dialing: The potential effect of sample characteristics on the conversion of nonresidential telephone numbers. Am J Epidemiol 136:1393–1399.
Waksberg J. 1978. Sampling methods for random digit dialing. J Am Statist Assn 73:4–46.
Ward EM, Kramer S, Meadows AT. 1984. The efficacy of random digit dialing in selecting matched controls for a case–control study of pediatric cancer. Am J Epidemiol 120:582–591.
Wartenberg D, Northridge M. 1991. Defining exposure in case–control studies: A new approach. Am J Epidemiol 133:1058–1071.
Wartenberg D, Rhoads GG, Northridge M. 1991. Cutpoint bias: Definition and Assessment Using Q–Q Plots (abstract). Am J Epidemiol 134:736.
Wartenberg D, Savitz D. 1993. Evaluating exposure cutpoint bias in epidemiologic studies of electric and magnetic fields. Bioelectromagnetics 14:237–245.
Wertheimer N, Leeper E. 1979. Electrical wiring configurations and childhood cancer. Am J Epidemiol 109:273–284.
Wertheimer N, Leeper E. 1980. RE: Electrical wiring configurations and childhood leukemia in Rhode Island (Letter). Am J Epidemiol 111:461–462.
Zaffanella L. 1993. Survey of Residential Magnetic Field Sources. Palo Alto, CA: Electric Power Research Institute. TR-102759 (2 vols.).
Zaffanella L, Greenland S, Iriye R. 1998. Application of the Case-Specular Method to the Los Angeles Study of Wire Codes and Childhood Leukemia. Palo Alto, CA: Electric Power Research Institute. TR-110686.
Zaffanella L, Greenland S, Savitz DA, Pearson R, Iriye R. 1997. Application of the Case-Specular Method to the Savitz Denver Study Residences. Palo Alto, CA: Electric Power Research Institute. TR-107751.
Zaffanella L, Savitz DA. 1995. Development of the Residential Case-Specular Epidemiologic Investigation Method. Palo Alto, CA: Electric Power Research Institute. TR-105900.
Zaffanella L, Savitz DA, Greenland S, Ebi KL. 1998. The residential case-specular method to study wire codes, magnetic fields, and disease. Epidemiology 9:16–20.

APPENDIX: DISCUSSION AT WORKSHOP

Stolwijk noted that worries about bias first appear as the design of a study is developed. It is important to perform a pilot study, especially for a very expensive study, since one can thus identify problems that could be very difficult to fix once the study has begun. He also reminded us that the most important and common study restraints—i.e. those of resources, available population and power, and exposure assessment and distribution of exposure—continue to test the ingenuity of epidemiologists. He concurred with Buffler that these restraints especially plague studies that investigate very low risks, relatively low exposures, or relatively rare diseases. Stolwijk also said that some studies which will be reporting their results soon have their own problems, as well. The British study, for instance, will not have the usual problems due to random-digit dialing, but it will have its own problems due to dissimilarity of the hospital registries employed. On the other hand, researchers find the well developed Swedish registries an enviable resource, but Swedish epidemiologists wish they had a larger population from which to draw subjects.

In general, Stolwijk continued, it is difficult to identify any sources of bias in a study and to analyze the influence of such biases. Usually published studies do not permit outside analysis of the effect of biases, and sometimes the studies do not even discuss possible sources of biases or why they might have appeared in the study. He went on to say that biases occur inevitably because researchers must make tradeoffs in the design stage of a study, even though they know that some of these design choices will produce problems later. It is impossible to meet all the constraints that face researchers and still avoid all possible biases that may appear. Stolwijk also agreed with Wartenberg that one cannot find a study free of all possible sources of bias, but one can luckily find a few studies where a reader can actually assess how strongly biases may have affected the outcome.

Wartenberg added that data sets are normally kept by the principal investigators and that there is no central repository where one can find the data and investigate questions of bias. If one wants to investigate bias in a study, one must contact each investigator and get clearance to obtain the data, an involved and time consuming process.

Bringing up selection bias in particular, Savitz related how in the late 1980s, Louise Brenton presented a very cursory slide of experience with random-digit dialing over the years, and he found that the effectiveness of random-digit dialing has decreased since it was first introduced. This trend, caused mostly by cultural factors, seems to be continuing; unfortunately, we do not have a better technique available for most general populations in the United States. Savitz also said that selection bias probably does really occur

and even though we don't know the magnitude of it, it must affect studies in important ways. If researchers systematically miss people who would have a certain wire code classification, he explained, the researchers would inevitably arrive at the wrong conclusions. Savitz added that he was very pessimistic about improving the present methods in the U.S. for selecting cases and controls. For example, if he were to repeat his Denver study, he is not confident that he could find controls that would illuminate the data from the original study.

Wartenberg suggested that using neighborhood controls might reduce selection bias. With advances in geographic information systems, he said, researchers can accurately map the area around a case residence and deduce important information about the population in that area. However, he added, this approach would introduce its own biases, such as over-matching. Langholtz replied that neighborhood controls have worked well in Los Angeles studies for a long time, and epidemiologists in that city have developed a very good system for finding and verifying neighborhood controls. He also said that he believes that researchers should use several different study designs to investigate a possible health risk, each of these studies inevitably having a different set of biases. The researchers can then compare the different studies and their results, informed by what they already know about the biases implicit in the study designs. Then an investigator would feel more comfortable with his or her conclusions if the different studies, with different biases, all yield similar results.

Ebi explained that in order to evaluate the hypothesis that wire codes are associated with a factor confounding the association with leukemia in certain neighborhoods, EPRI sponsored a study using the case-specular method. This method has a variety of assumptions, including the assumption that the distribution of wire codes in the controls and the distribution of wire codes in the speculars for those controls is symmetric. If that assumption is not met, an asymmetry would reveal control selection bias. They found that the control-specular matrices were not symmetric in Denver or LA, and they were clearly not symmetric in opposite ways, although the reason for this remained unclear. Wertheimer related that she did what was essentially a case-specular study for the California EMF Project's adult study of the Denver area sample. By using an actual house, randomly distributed within an eight block region around the case, they avoided the problem of a biased, fictional house.

Neutra noted that the proportion of VHCC in all these studies was down around 3%, a small number. However, when Pearson and Wachtel used a computerized method to wire code a large citywide sample, the proportion of VHCC homes became a little larger, approximately 7%. Neutra wondered if the difference, which is not very large, implies that there was selection bias in these studies, or if the difference is simply within the computer's margin of error. He also argued that in Wertheimer's case the researchers avoided selection bias by selecting controls from birth certificates. Wertheimer corrected this statement, reminding Neutra that in her childhood study they did not use the VHCC classification, so she could not say whether her prevalence of VHCC controls in that study was similar to Savitz's or Pearson's.

Ebi clarified that when GIS methods were used to automatically wire code Denver, the exercise was restricted to neighborhoods that existed when Savitz's study was conducted, and it only included census blocks that actually had children living in them. Furthermore, the case-specular study found a reasonably good correlation between the wire code assigned by direct observation and that assigned by the computer methods, but it did not find an exceptionally high correlation. However, there was one major unresolved issue: the power system had changed in undetermined ways between studies, and the researchers did not know how to account for those changes.

Wertheimer remarked that if distances were measured from the center of the roof in the computerized method, this would result in more VHCCs being found than would be the case if the original Wertheimer and Leeper protocol were used.

Neutra argued that if selection bias in the controls had led to a positive result in the London study, then we should be able to compare Preston-Martin's brain cancer cases to London's controls and find an effect. However, when we make that comparison, we do not find any effect. Similarly, researchers did not find a strong association in LA between mobility and wire code when they used the data reported by Jones.

Tarone expressed his conviction that bias has a greater effect on the quality of a study than does confounding. Because of the increasing difficulty of finding controls, he wondered if investigators are hurting themselves by trying to match too closely on too many things. Perhaps, he said, it would be better to match less closely and stress other things instead. Carpenter responded that there is a consensus that a couple of variables generally need to be matched. Originally, epidemiologists thought that matching successfully controlled for confounding. However, he said, on closer examination, it's not clear that matching does an adequate job. Even when there is confounding in a study, it's never clear whether confounding or bias is the biggest problem.

Savitz observed that geographic matching poses the most difficulties and that researchers have two options. On one hand, they can take all the cases from a given area of a city, such as Denver or Los Angeles, and try to create a random sample of the population by doing their best with RDD. On the other hand, they can change the last few digits of the case's phone number, in which case they are performing a variant of individual matching. Researchers should choose a method depending on how densely phoned the area under consideration is, which in turn affects how tightly matched the subjects will be on other variables. In Denver it appeared that the second approach would keep much of the meaningful variability in the wirecode. Then Savitz noted that in the Denver inner city area a single telephone exchange forms a tighter geographic group than it would in a more remote area, but he still did not think geographic proximity is one of the most important problems he confronted. Carpenter replied that the importance of geographic density depends on the geographical area under study. In some areas of suburban New Jersey, the phone number defines a more homogeneous region than a zip code in the same region. DelPizzo added that as the demographics of an area become more homogeneous, the wire codes become more homogenous as well. If a case is found in a new development, and all the phone numbers in the area have the same initial digits, it is very likely that all the possible controls have underground power as well. Moreover, over the past decade or so, the phone line density has become higher than the phone line density of households, now that many families have more than one line.

Langholtz agreed that we probably do not need close matching as much as we used to think, but he thought we should pay more attention to defining the study base. He wondered how we can begin to better define the demographics of the study base and their origin.