Residential Magnetic Fields and Childhood Leukemia: A Meta-Analysis

Daniel Wartenberg, PhD

ABSTRACT

Objectives. This article uses meta-analysis methodology to examine the statistical consistency and importance of random variation among results of epidemiologic studies of residential magnetic field exposure and childhood leukemia.

Methods. A variety of meta-analytic statistical methods were applied to all available studies combined and on subgroups of studies chosen by exposure characteristics. Sample sizes and fail-safe n's were calculated to determine the robustness of results and the potential role of publication bias.

Results. Most studies show elevated but not statistically significant odds ratios. Results for exposures assessed by wire codes, distance, and/or historically reconstructed fields are relatively consistent, homogeneous, and positive, while those for direct magnetic field measurements are consistent, homogeneous, and marginally protective. Several unpublished studies, or a single unpublished study with several hundred subjects, would be needed to nullify the observed data.

Conclusions. The observed results identify a consistent risk that cannot be explained by random variation. The data supporting magnetic fields as the principal risk factor are suggestive but inconsistent. Additional studies using innovative designs that focus on highly exposed children offer the most hope of untangling this issue. (*Am J Public Health.* 1998;88:1787–1794)

Since the publication of a seminal study by Wertheimer and Leeper,[1] scientists have attempted to make sense of provocative and sometimes conflicting studies about the possible association between exposure to electric and magnetic fields and the incidence of disease. As this controversy continues, organizing and reviewing the extant data can provide important insights into the consistency of the results, gaps in our investigative strategies, and limitations in our understanding. Toward that end, this article presents a meta-analysis of the most compelling subset of these data: data on residential exposure to magnetic fields and the incidence of childhood leukemia. It is an attempt to gain an understanding of the importance of individual studies in driving overall conclusions about a possible link between magnetic fields and cancer and of the constraints that would be necessary on any future study for it to have sufficient statistical power to influence the present overall conclusions.

Meta-analysis is a statistical method used to provide a single summary risk estimate based on a set of similar epidemiologic studies.[2,3] It is applied most often to clinical trial data in which the major differences among studies are the specific populations examined rather than characteristics of the study designs. The validity of broadening the application of this method to environmental epidemiology has led to controversy because of the heterogeneity in results that often arises from design differences among studies in exposure assessment, confounder assessment, subject selection, and so forth.[4–8] However, meta-analysis methods can also be used in a less statistically rigorous manner to evaluate the strength, consistency, and robustness of an exposure–disease relationship. This article presents one such application.

Of the 16 epidemiologic studies to date (see Table 1), some have reported positive results and others have found no association. Scientists disagree about the quality, bias, accuracy, uncertainties, and many of the statistical analyses in each of these studies; thus, there are differing interpretations of the likelihood of a possible association overall.[9–11] While some investigators question the validity of drawing inferences based on 16 or fewer studies with apparently inconsistent results, the ubiquitous nature of exposure to magnetic fields from power lines makes even a weak association a public health issue of substantial concern. Meta-analysis, while no better than the data on which it is based, can help frame the public health debate while also providing insights for the design of additional studies.

Five sets of investigators have previously conducted meta-analyses of childhood cancer and residential exposure to magnetic fields. A report by Great Britain's Advisory Group on Non-Ionising Radiation of the National Radiological Protection Board summarized the results of the residential exposure studies, providing pooled odds ratio estimates for 3 studies for each exposure metric.[11] For wire codes (a categorical exposure rating scheme based on wire size and distance from the residence), excluding the Wertheimer and Leeper study,[1] the group found an elevated and statistically significant odds ratio. For distance from source of magnetic fields and for measured magnetic fields, the pooled odds ratios were elevated but not statistically significant.

Ahlbom et al.[9] combined results from 3 recent studies conducted in Nordic coun-

The author is with the Department of Environmental and Community Medicine and the Environmental and Occupational Health Sciences Institute, University of Medicine and Dentistry of New Jersey–Robert Wood Johnson Medical School, Piscataway, NJ.

Requests for reprints should be sent to Daniel Wartenberg, PhD, Environmental and Occupational Health Sciences Institute, 170 Frelinghuysen Rd, Piscataway, NJ 08855 (e-mail: dew@eohsi.rutgers.edu).

This article was accepted June 25, 1998.

TABLE 1—Epidemiologic Studies of the Association Between Exposure to Electric and Magnetic Fields and Childhood Cancer

Study	Design	Variables Matched on	Potential Confounders Assessed	Exposure: Wire Codes	Distance	Spot Measurements	24-Hour Measurements	Calculation of Historical Magnetic Fields
Wertheimer and Leeper[1]	Case–control	Age Residence location	Sex Onset age SES Family pattern Traffic density	X				
Fulton et al.[33]	Case–control	Age	SES Onset age	X				
Tomenius[39]	Case–control	Age Sex Residence location	Residency permanence			X		
McDowall[21,a]	Cohort	None	. . .		X			
Savitz et al.[34]	Case–control	Age Sex Residence location	SES Traffic density Maternal age Maternal smoking in pregnancy Paternal education	X		X		
Coleman et al.[36]	Case–control	Age Sex Residence location Onset age	Age Sex SES		X			
Lin and Lu[23,b]	Case–control	Age Sex	None		X			
Myers et al.[19,c]	Case–control	Age Sex Residence location	Residence type		X			
London et al.[35]	Case–control	Age Sex Ethnicity	SES Residence type Appliance use Parental occupation Environmental exposures	X		X	X	
Lowenthal et al.[20,d]	Case–control	Unknown	. . .		X			
Olsen et al.[12]	Nested case–control	Age Sex	Onset age					X
Feychting and Ahlbom[14]	Nested case–control	Age Sex Residence location	SES Residence type Diagnosis year Nitrogen dioxide exposure		X	X		X
Fajardo-Gutierrez et al.[37]	Case–control	Unknown	. . .		X			
Petridou et al.[38]	Case–control	Unknown	. . .		X			
Verkasalo et al.[13,]	Cohort	5-y age group	. . .					X
Schreiber et al.[22,a]	Cohort	Residence date Dutch nationality	. . .		X			

Note. SES = socioeconomic status.
[a]Excluded from the meta-analysis because subjects included both children and adults.
[b]Excluded from the meta-analysis because data were unpublished.
[c]Excluded from the meta-analysis because data were incomplete.
[d]Excluded from the meta-analysis because the study's analysis was incomplete.

tries,[12–14] arguing that they were more similar to one another than to other studies (each used a population registry and estimates of historical exposure) and were thus appropriate for use in a meta-analysis. By combining risk ratios found in these studies with weights proportional to the inverse of the variance, Ahlbom et al. found statistically significantly elevated risk ratios for leukemia.[15]

Washburn et al. conducted a set of meta-analyses for leukemia, lymphoma, and nervous system cancers.[16] They found that the combined risks estimated from 13 studies were elevated for all 3 outcomes; those for leukemia were statistically significantly elevated for several alternative exposure groupings.

In their meta-analysis, Miller and colleagues compared exposure measurement techniques.[17] Using 7 studies to assess the childhood leukemia risk, they found a statistically significant positive association of childhood leukemia with wire codes (n = 4), distance (n = 2), and calculated index (n = 1), and a nonsignificant positive association with spot measures (n = 4).

Meinert and Michaelis conducted the most recent meta-analysis. They considered reported cancers by type, for a variety of exposure metrics, grouping 2 to 6 studies at a time.[18] They found statistically significant associations of leukemia with wire codes and magnetic fields but not distance, and nonsignificant positive associations for lymphomas, central nervous system tumors, and all tumors (data for the highest magnetic field were statistically significant).

The present study enhances previous meta-analyses in 2 ways. First, it assesses heterogeneity among studies, influence of individual studies, and possible publication bias. Second, it considers all studies to date, both stratified by exposure type and combined.

Methods

There are 2 important methodological aspects to conducting a traditional meta-analysis: study selection and choice of statistical methods of summarizing results across studies.

Selecting the Studies

Studies for this analysis were identified from previous reviews and by asking researchers active in this field for recommendations. Overall, 16 studies of residential magnetic field exposures and childhood cancer were reviewed (see Table 1). Of these, 5 were excluded because data presentation or analysis was incomplete,[19,20] children were not analyzed independently of older subjects,[21,22] or the data were not published and thus were inaccessible.[23] In the remaining 11 studies, variation is introduced by outcome studied (1 was a mortality study, the rest incidence studies); source of data (hospital records, incidence registry, birth registry, death registry); maximum age of subjects (from 10 to 20 years); and exposure metric used (wire codes, distance from electrical source, measured magnetic field, and historical reconstruction of magnetic field).

The data were stratified by exposure metric, and (relatively) consistent exposure cutpoints (i.e., dichotomous exposure classifications) were used in each analysis to maximize the consistency of exposure among subjects across studies.[5] Although Washburn et al.[16] reported that use of alternative cutpoints did not alter their meta-analysis, Wartenberg and Savitz[24] showed for an individual study and Meinert and Michaelis[18] showed for meta-analysis that cutpoint choice can be important. In this article, separate analyses are reported for (1) studies using wire codes, considering 2 alternative exposure cutpoints; (2) studies using only distance from electrical source, considering 2 distance cutpoints; (3) studies using wire codes or distance, considering 3 alternative cutpoints; (4) studies using magnetic field measurements, considering 1 cutpoint; (5) studies using historical reconstruction of magnetic fields, considering 1 cutpoint; and (6) all studies together. Three separate analyses were conducted for the last category, using (a) the exposure type/cutpoint that gave the smallest probability of the null hypothesis, (b) the exposure type/cutpoint that gave the largest probability, and (c) the highest exposure proportion in each study.

Selecting the Statistical Methods

A variety of methods have been used to assess combined effects, to identify heterogeneity, and to conduct influence analyses.[2,3,25–27]

The simplest method, called *vote counting*, results in tallies of the number of studies with positive results, the number with negative results, and the number with null results. Many criticize this approach because it has low statistical power and because the summary measure does not incorporate the observed effect size or sample size.[25] However, even with low power, positive results can be interpretable and thus are reported here, although merely for guidance.

The *combined probability test* combines the logarithms of individual study probabilities (P values) into a χ^2 distributed statistic, P, with the degrees of freedom equal to twice the number of studies combined.

Statistics that summarize the individual effect sizes use either of 2 statistical models: *fixed effects* or *random effects*. The fixed-effects model assumes that the studies have the same true effect size. Within-study precision (i.e., an overall treatment effect) is assessed by weighting individual study results by the inverse of the variance. The random-effects model assumes that the studies included have different true effect sizes that form a statistical distribution, and it estimates the average effect for the whole population on the basis of observed data, including both interstudy variation (i.e., a sampling effect) and intrastudy precision (i.e., a treatment effect).[28,29] Model choice may be based on results of the Q test for heterogeneity, which assesses constancy of treatment effects.[6,30]

Influence analysis is the recalculation of summary indices for a set of studies, leaving out one study at a time. It indicates the importance of each individual study in the combined summary statistic and enables one to determine whether any of the studies has a disproportionate influence.[8]

Publication bias can be assessed by combining z scores of individual studies and determining the number of additional null studies needed to reduce a statistically significant combined effect to nonsignificance.[31,32] This number is called the *fail-safe n.*

Similarly, using the combined-effect measure, one can assess how large a study would be required to balance the average of reported results if they were due to random fluctuations. This can be viewed as the sensitivity of the cumulative results, either to publication bias or to results of future studies. We can determine what size a single study would have to be to give a null summary statistic (i.e., an odds ratio of 1.0) if that hypothetical study had equal numbers of cases and controls, an exposure prevalence equal to that observed in reported studies, and an odds ratio equal to the reciprocal of the combined effect. Unlike the fail-safe n, this calculation uses the size of the effect measure, weights each study result by the inverse of its variance, hypothesizes a study with a protective rather than a null effect, and seeks a null rather than a nonsignificant combined effect.

Results

The 16 studies of residential exposure and childhood cancer used in the meta-analysis are described in Table 1. A table of the data used is available from the National Auxiliary Publications Service (NAPS; see Acknowledgments). Results of 2 particular analyses are shown in Tables 2 and 3. Summaries of all analyses are provided in Table 4.

Results of analyses of data from individual studies and selected meta-analyses are shown in Figure 1. The odds ratio (dot) and its 95% confidence interval (vertical line) are plotted for each dichotomous cutpoint of each exposure metric of each study. The expected null effect, an odds ratio of 1.0, is shown by the horizontal line. The ordinate scale is logarithmic. The data are arranged by exposure metric used; within exposure metric, ordered by the value of that metric; and for metrics of the same value, ordered by the year of the study.

For instance, the first 3 groups of studies are wire code studies: very low current configuration (VLCC), low current configuration (LCC), and high current configuration (HCC), each as reported by the authors. (The VLCC category includes buried, underground, and end-pole configurations.) The number below each line identifies the specific data as listed in the table available from NAPS, by the plot ID number in the last column of that table. The first data line, denoted by 1, is from Wertheimer and Leeper,[1] for wire codes at residence at the time of birth with end poles as the exposure cutpoint; the second, denoted by 4, is from the same study but assessed at residence at the time of death, with the same exposure cutpoint. The next line, denoted by 7, is from Fulton et al.,[33] with the reported cutpoint of VLCC.

The next 2 sets of data used distance as the exposure metric. The first set used distance of distribution lines and the second set used distance of other features (e.g., transformers, substations, and transmission lines in Fajardo-Gutierrez et al.[37]).

The final 3 sets of data represent studies in which magnetic fields were reported. The first set reported spot measurements; the second, 24-hour (day-long) measurements; and the third, historical calculations. Meta-analysis results for data subsets are denoted in the fig-

TABLE 2—Results of Analysis of Studies Using Wire Codes (Low Current Configuration)

Study	No. of Exposed Cases	No. of Exposed Cases Expected	Influence Analysis[a] Combined *P*	Individual OR (95% CI)	$OR_{Fixed\ effects}$ (95% CI)	$Pr\{Q_{het}\}$[b]	$OR_{Random\ effects}$ (95% CI)
Wertheimer and Leeper[1,c]	52	22.77	.01	2.28 (1.34, 3.91)	1.35 (1.06, 1.73)	.16	1.36 (0.97, 1.91)
Fulton et al.[33]	87	86.50	.00	1.00 (0.67, 1.49)	1.78 (1.36, 2.33)	.55	1.78 (1.36, 2.33)
Savitz et al.[34]	27	17.58	.00	1.54 (0.90, 2.63)	1.47 (1.15, 1.88)	.04	1.53 (0.97, 2.42)
London et al.[35]	122	72.46	.00	1.68 (1.14, 2.48)	1.39 (1.05, 1.82)	.05	1.48 (0.91, 2.42)
All combined	288	199.31	.00	. . .	1.48 (1.18, 1.85)	.08	1.52 (1.08, 2.14)

Note. Sample size needed to balance observed results = 1294; fail-safe n = 16. OR = odds ratio; CI = confidence interval.
[a]Influence analysis is the recalculation of summary indices for a set of studies, leaving out one study at a time.
[b]Probability of homogeneous study results, based on the χ^2 heterogeneity test statistic Q.[6,28,30]
[c]Data used in the analysis reported here were for residential exposure at the time of birth.

ure by numbers preceded by an M. These are listed in the plot ID column in Table 4. One striking observation is the preponderance of odds ratios at or above the null effect line. Only 8 of the 53 odds ratio dots fall below the null effect line, and 4 of these are for measured magnetic fields. This figure is analogous to an unweighted vote-counting assessment and strongly suggests an association between exposure to magnetic fields and childhood leukemia.

Wire Code Studies

I evaluated the 4 wire code studies using 2 different groupings of several categories, using LCC and then HCC as the dichotomous cutpoint.[1,33–35] The results were fairly similar, although, surprisingly, the results showed slightly stronger evidence for an association in the LCC analysis than in the HCC analysis.

Distance Measurement Studies

Four studies used distance to estimate exposure, with cutpoints of 50, 100, and 200 m.[14,36–38] The proportion of exposed cases was similar with both cutpoints, but this was due largely to the inclusion of the Petridou et al. study in the 50-m class.[38] Exclusion of this study brought the proportion exposed in the 50-m-cutpoint category down to 10%, substantially less than the 27% exposed in the 100-m-cutpoint category. The results for the higher exposure category (i.e., 50 m) showed stronger positive associations, as would be expected if an association exists.

Studies Using Wire Codes and Distance Measurements

Since distance is an important component of the wire coding scheme, this analysis combined the following metrics: LCC with 100 m, LCC with 50 m, and HCC with 50 m. For studies with both distance and wire code data, wire codes were used. Seven studies were available. All 3 combinations gave fairly similar results, showing statistically significantly elevated odds ratios, moderate insensitivity to deletions of single studies, and a need for large fail-safe n's and sample sizes to balance the observed data, indicating substantial robustness to results from additional studies.

Spot Measurement Studies

Spot measurements were combined from 4 studies, by using the cutpoint closest to 2.0 mG across all studies.[14,34,35,39] For the Savitz et al. study,[34] exposures measured with the appliances turned off (low power) were used. A nonsignificant, slightly negative result was found.

Studies Using Calculations of Historical Magnetic Fields

Three Nordic studies reported calculations of historical magnetic fields.[12–14] Using a cutpoint close to 0.2 μT resulted in a positive, statistically significant result.

All Studies Combined

An analysis of all studies combined, using the smallest probability, which biases the result, showed significance for both heterogeneity and elevated odds ratio. The single most influential study was that of Fulton,[33] which pulled down the odds ratio from 1.81 to 1.64. When the data using the highest exposure or cutpoint were combined, the analysis showed marginal heterogeneity and a statistically significant, elevated odds ratio. A combined-studies analysis using the largest probabilities, which biases the result toward protection, showed homogeneity, and the slightly positive effect was not statistically significant. The deviation of the odds ratio from the null was far less than that observed for the combined-studies analysis using the smallest probabilities.

Discussion

This paper evaluates the role of random variation in epidemiologic studies examining residential magnetic field exposure and childhood leukemia. The results of this meta-analysis suggest that the data cannot be explained statistically on the basis of random fluctuations alone, and these results show little sensitivity to how the data were grouped. Publication bias is an unlikely explanation, inasmuch as several negative studies have been published and, generally, a large fail-safe n or study size is needed to negate the observed results. However, the inconsistency of statistical results when measured magnetic fields are used, vs other exposure metrics, remains an enigma. Exposure misclassification, which typically biases the results toward the null, may play a role.

Consider the evidence piece by piece. First, and perhaps most striking, is the simple graph of the odds ratios (see Figure 1). The preponderance of positive associations across metrics (except spot measures) and exposure cutpoints speaks to the overall consistency of the data. While there are issues of independence of results, weighting of multiple results from the same study, and possible biases in each study, the pattern is not random.

Further, the vote-counting results show that in 11 of the 12 groupings, at least half of the results had elevated odds ratios, whereas this pattern was expected in only 6 groupings (Table 4). The exception is for the grouping of the weakest (or most protective) responses in each study. In 9 of the 12 groupings, the number of statistically significant odds ratios exceeded the 5% expected by chance, whereas this was expected only once. In 8 of these 9 groupings, at least 25% of the results were statistically significant, and in 5 groupings more than 30% were statistically significant. The analyses that had no statistically significant results were those for 1 group of distance studies, the group of spot magnetic

TABLE 3—Results of Analysis of Studies Using Wire Codes (Low Current Configuration [LCC]) and Distance

Study	Exposure Definition	No. of Exposed Cases	No. of Exposed Cases Expected	Influence Analysis[a] Combined *P*	Individual OR (95% CI)	$OR_{Fixed\ effects}$ (95% CI)	$Pr\{Q_{het}\}$[b]	$OR_{Random\ effects}$ (95% CI)
Wertheimer and Leeper[1]	LCC (at birth)	52	22.77	.01	2.28 (1.34, 3.91)	1.31 (1.06, 1.62)	.29	1.32 (1.03, 1.68)
Fulton et al.[33]	LCC	87	86.50	<.01	1.00 (0.67, 1.49)	1.58 (1.26, 1.99)	.32	1.58 (1.23, 2.03)
Savitz et al.[34]	LCC	27	17.58	<.01	1.54 (0.90, 2.63)	1.40 (1.13, 1.72)	.09	1.43 (1.04, 1.96)
Coleman et al.[36]	Distance <100 m	36	38.77	<.01	0.93 (0.54, 1.60)	1.51 (1.22, 1.86)	.21	1.54 (1.18, 2.01)
London et al.[35]	LCC	122	72.46	<.01	1.68 (1.14, 2.48)	1.33 (1.06, 1.67)	.12	1.39 (1.01, 1.91)
Feychting and Ahlbom[14]	Distance <100 m	12	7.42	<.01	1.62 (0.79, 3.30)	1.40 (1.14, 1.72)	.09	1.42 (1.05, 1.92)
Fajardo-Gutierrez et al.[37]	Distance <200 m	6	2.61	<.01	2.30 (0.54, 9.80)	1.40 (1.15, 1.71)	.10	1.42 (1.07, 1.87)
All combined		342	248.11	<.01	. . .	1.41 (1.16, 1.72)	.14	1.44 (1.10, 1.87)

Note. Sample size needed to balance observed results = 1765; fail-safe n = 26. OR = odds ratio; CI = confidence interval.
[a]Influence analysis is the recalculation of summary indices for a set of studies, leaving out one study at a time.
[b]Probability of homogeneous study results, based on the χ^2 heterogeneity test statistic Q.[6,28,30]

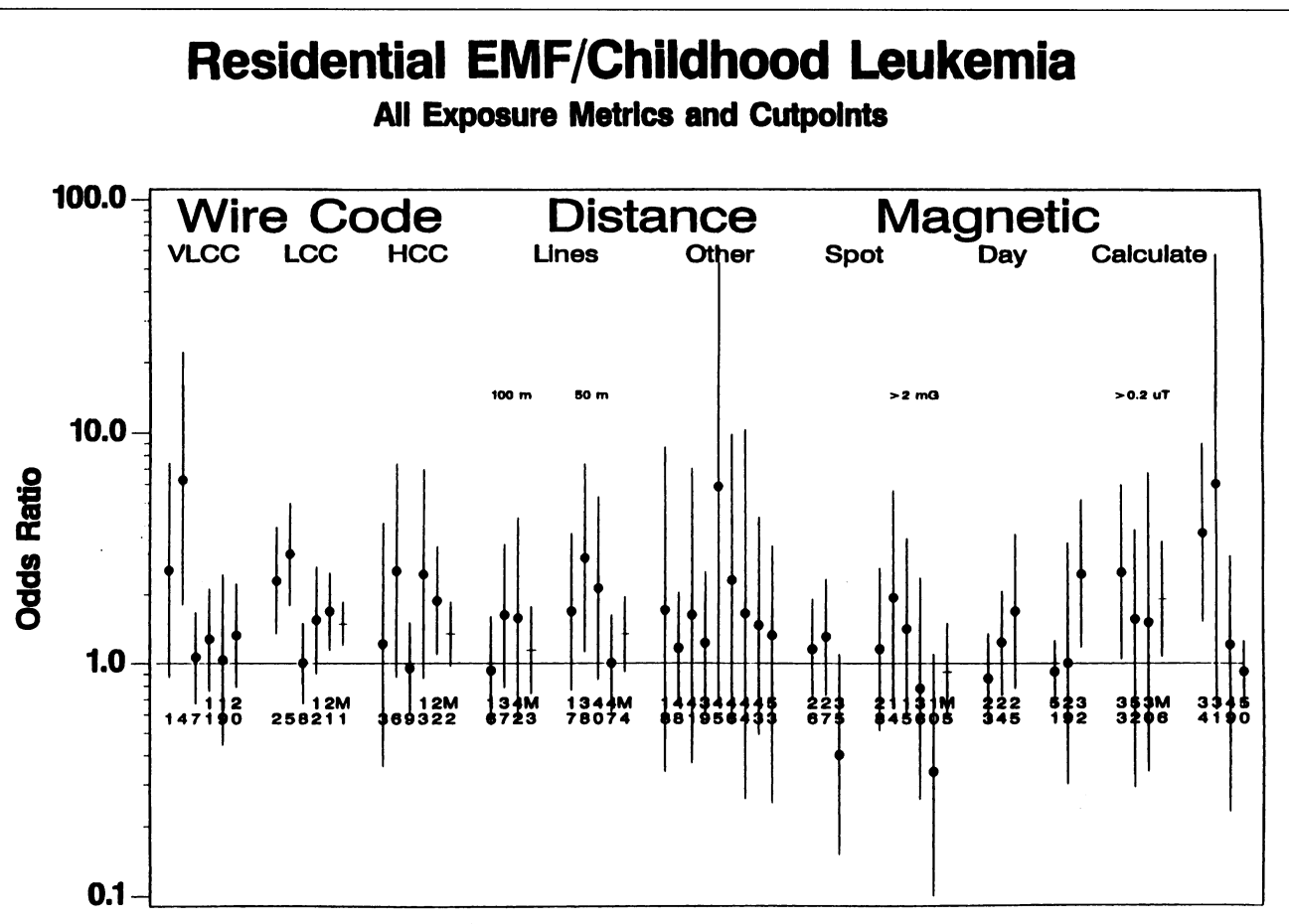

FIGURE 1—Results of analyses of data from individual studies of the association between exposure to electric and magnetic fields and childhood cancer and selected meta-analyses (see Results section).

field measurements, and the group of studies using the largest *P* values. The results of the combined probabilities test are similar to those for the vote counting.

The results of the combined odds ratio assessments, performed with both the fixed-effects and random-effects models, show a similar pattern. For 10 of the 12 groupings, odds ratios were elevated, ranging between 1.14 and 1.90; for 7 of the 12, the 95% confidence intervals excluded 1.0. These results were not sensitive to the deletion of a single study. The exposure groups that did not show an elevated risk were the groups using the spot magnetic field data and the group using the largest probability from each study.

One of the most important uses of the random-effects model is to explore study differences when heterogeneity is present.[6] The combinations that showed statistically significant heterogeneity were the grouping of all studies using the smallest probabilities and the grouping of all studies using the highest exposures. This effect may be due, in part, to the different exposure metrics used across studies, which likely reflect different aspects of magnetic field exposure. When stratified

TABLE 4—Summary of Alternative Meta-Analyses

Data Set and Cutpoint	No. of Studies	Proportion Exposed	Vote Counting		Combined *P* Values		Combined $OR_{Fixed\ effects}$			Combined $OR_{Random\ effects}$			
			No. Positive (%)	No. Statistically Significant (%)	$Pr\ (P)^a$	Fail-Safe n	All Studies OR (95% CI)	Influence Analysis[b] Range of ORs	Size Needed	$Pr\ (Q_{het})^c$	All Studies	Influence Analysis[b] Range of ORs	Plot ID
Wire codes (LCC)	4	0.40	3 (75)	2 (50)	<.01	16	1.48 (1.18, 1.85)	(1.35, 1.78)	1294	.08	1.52 (1.08, 2.14)	(1.36, 1.78)	M1
Wire codes (HCC)	4	0.13	3 (75)	1 (25)	.02	2	1.34 (0.97, 1.85)	(1.12, 1.85)	1319	.18	1.42 (0.90, 2.24)	(1.21, 1.85)	M2
Distance (<100 m)	3	0.27	2 (67)	0 (0)	.13	...	1.21 (0.80, 1.83)	(1.04, 1.73)	460	.32	1.24 (0.78, 1.96)	(1.13, 1.73)	M3
Distance (<50 m)	3	0.26	2 (67)	1 (33)	.02	2	1.34 (0.92, 1.95)	(1.15, 2.09)	564	.12	1.52 (0.83, 2.78)	(1.18, 2.09)	M4
Wire codes (LCC) and distance (<100 m)	7	0.35	6 (86)	2 (29)	<.01	26	1.41 (1.16, 1.72)	(1.31, 1.58)	1765	.14	1.44 (1.10, 1.87)	(1.32, 1.58)	...
Wire codes (LCC) and distance (<50 m)	7	0.34	5 (71)	3 (43)	<.01	36	1.44 (1.19, 1.75)	(1.35, 1.61)	1875	.08	1.50 (1.14, 1.97)	(1.38, 1.63)	...
Wire codes (HCC) and distance (<50 m)	7	0.19	5 (71)	2 (29)	<.01	14	1.34 (1.05, 1.71)	(1.23, 1.53)	1696	.16	1.42 (1.03, 1.97)	(1.33, 1.58)	...
Spot measurements (≥2.0 mG)	4	0.10	2 (50)	0 (0)	.52	...	0.95 (0.57, 1.57)	(0.78, 1.20)	653	.17	0.92 (0.47, 1.78)	(0.74, 1.20)	M5
Calculated historic fields (≥0.2 μT)	3	0.02	3 (100)	1 (33)	.03	2	1.90 (1.07, 3.39)	(1.54, 2.19)	2654	.72	1.90 (1.07, 3.39)	(1.54, 2.19)	M6
All combined													
Smallest *P*	11	0.17	10 (91)	3 (27)	<.01	66	1.54 (1.25, 1.91)	(1.48, 1.71)	2460	.02	1.67 (1.18, 2.36)	(1.53, 1.81)	...
Largest *P*	11	0.15	5 (45)	0 (0)	.52	...	0.91 (0.76, 1.08)	(0.89, 0.93)	3839	.72	0.91 (0.76, 1.08)	(0.89, 0.93)	...
Highest exposure	11	0.05	9 (82)	1 (9)	.05	15	1.31 (1.01, 1.69)	(1.19, 1.52)	4893	.10	1.42 (0.98, 2.04)	(1.23, 1.55)	...

Note. OR = odds ratio; CI = confidence interval.
[a]Probability test that combines the logarithms of individual study *P* values.
[b]Influence analysis is the recalculation of summary indices for a set of studies, leaving out one study at a time.
[c]Probability of homogeneous study results, based on the χ^2 heterogeneity test statistic Q.[6,28,30]

by exposure metric, none of the combinations showed statistically significant heterogeneity, suggesting that statistics summarizing all studies may be misleading.

Certain studies showed up repeatedly as responsible for a substantial portion of the heterogeneity. The Fulton et al. study had the largest effect on analyses of studies using wire codes or wire codes and distance.[33] The Tomenius study had the largest effect on analyses of studies using spot measures or all studies combined.[39] The study of Fulton et al. has been criticized for selection bias,[40] and that of Tomenius for exposure misclassification.

Publication bias was assessed by the fail-safe n. The fail-safe n suggests that for the observed excess to be due to publication bias, there would have to have been at least a dozen unpublished, negative studies, except for those groupings consisting of very few studies. In view of the strong interest in this topic among scientists, it seems unlikely that any investigator would have trouble getting even a negative study published. Indeed, 2 of the 11 published studies reported protective results (i.e., odds ratios less than 1.0), and 8 of 11 did not have probabilities less than .05 for any exposure cutpoints, suggesting that negative and nonsignificant results are readily publishable.

The sample size needed to negate positive results is a measure of overall robustness. Generally, a single study of more than 1000 subjects (500 cases) would be required to balance the results of any of the meta-analyses. To balance the results of the meta-analysis of all studies combined, a sample of more than 2000 subjects (1000 cases) would be required. While some may view this assessment as artificial (i.e., requiring a study to show a protective rather than null effect), the sample size needed is useful in assessing the likely impact of another study on the combined results.

Two contradictions remain: (1) Why do the spot measures of magnetic field strength not show an association with leukemia, even though all the other exposure metrics do? and (2) Why do the data not show a consistent dose–response relationship? With respect to the spot measures, consider 3 explanations. First, other exposure metrics may be markers for the true risk factor, and the true risk factor may not be related to magnetic field strength (sometimes called factor X). A number of plausible such risk factors have been investigated, but none have explained the association. Second, other metrics may be more biologically relevant measures of exposure than spot measures. They may be more representative of long-term, integrated averages of magnetic

field strength or of some other aspect of magnetic field exposure that causes disease, for example, peak field strength, field variability, or time above a specific threshold. Third, the relatively limited proportion of individuals for whom direct measurements of magnetic fields were obtained may have resulted in bias (e.g., in the Savitz et al. study, the rate was 64% [35% for cases][34]; in the London et al. study, it was 54% [60% for cases][35]; and in the study by Feychting and Ahlbom, it was 62% [62% for cases][14]).

The inconsistency with regard to dose–response relationship has not been assessed in a statistically rigorous manner. A comparison of alternative cutpoints must be evaluated more carefully than by simply comparing odds ratios.[24] Possible explanations for the observed distribution include misclassification, low spot-measurement rates, and unaccounted-for geographic variation in wire codes. Additional analyses using more rigorous statistical methods (e.g., splines) may help to define the nature and extent of the inconsistencies.

While this paper was in review, 4 additional reports were published on studies of the possible relationship between exposure to magnetic fields and childhood leukemia.[41–44] The US study is the largest case–control study on this topic conducted to date, with more than 600 cases and 600 controls; that study showed no association of cancer risk with wire codes but a positive association with measured magnetic fields.[41,45–50] The German report, with data pooled from 2 studies in different parts of Germany, showed an association of cancer risk with measured magnetic fields.[42,51] The Greek study did not show an association of cancer risk with any of 4 metrics using various functions of voltage and distance from lines, nor with modified wire codes.[43] The Norwegian study, a nested case–control study using calculated residential magnetic field exposures, found no association between leukemia risk and time-weighted average exposure.

When the results of these studies are included in the meta-analysis, only small changes are found. The wire code/distance and calculated-exposure results moved closer to the null, largely owing to the results of Linet et al.[41] and Tynes and Haldorsen.[44] The spot-measure results moved from slightly protective (i.e., an odds ratio less than 1.0) to slightly risky (i.e., an odds ratio slightly greater than 1.0), also owing to the results of Linet et al.[41] The results for the combined grouping barely changed, reflecting the insensitivity of the results to the inclusion of new studies (unless they are markedly disparate from previous studies); this insensitivity is also reflected in the calculations for publication bias and needed sample sizes. None of the results changed in terms of statistical significance, and the likelihood of publication bias for the positive results decreased owing to the large size of these additional studies.

Overall, the data provide relatively strong and consistent support for a somewhat weak elevated risk of leukemia for children living in proximity to power lines. Another case–control study would likely be very expensive and only marginally informative. Our understanding of this issue would best be advanced by studies that address the inconsistencies. A study of high-exposure individuals, such as those living close to high-voltage transmission lines or in homes with high current configuration, might provide the most informative approach, increasing the exposure gradient and thus the statistical efficiency of the design. □

Acknowledgments

This project was supported in part by National Institute of Environmental Health Sciences Center grant ES05022.

I thank David Savitz and Nancy Wertheimer for providing more detailed exposure data than were available in the paper published by Wertheimer and Leeper in 1979. David Savitz, John Tukey, Daniel Driscoll, Richard Garwin, and Larry Toburen provided very useful advice during the development of the paper.

This meta-analysis was conducted originally for the National Academy of Sciences Committee on the Possible Biological Effects of Exposure to Magnetic Fields.[52]

See NAPS document no. 05483 for 2 pages of supplementary material. Order from NAPS in care of Microfiche Publications, 248 Hempstead Turnpike, West Hempstead, NY 11552-2664. Remit with your order a check or money order for $15.00 (US funds only, payable to Microfiche Publications) for photocopies or $5.00 for microfiche. For orders to be mailed outside the United States and Canada, add postage of $4.50 for photocopies or $1.50 for microfiche.

References

1. Wertheimer N, Leeper E. Electrical wiring configurations and childhood cancer. *Am J Epidemiol.* 1979;109:273–284.
2. Dickersin K, Berlin J. Meta-analysis: state of the science. *Epidemiol Rev.* 1992;14:154–176.
3. Petitti D. *Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis.* New York, NY: Oxford University Press; 1994.
4. Blair A, Burg J, Foran J, et al. Guidelines for application of meta-analysis in environmental epidemiology. *Regul Toxicol Pharmacol.* 1995;22:189–197.
5. Shapiro S. Meta-analysis/shmeta-analysis. *Am J Epidemiol.* 1994;140:771–778.
6. Greenland S. Invited commentary: a critical look at some popular meta-analytic methods. *Am J Epidemiol.* 1994;140:290–296.
7. Greenland S. Can meta-analysis be salvaged? *Am J Epidemiol.* 1994;140:783–787.
8. Olkin I. Re: "A critical look at some popular meta-analytic methods." *Am J Epidemiol.* 1994;140:297–299.
9. Ahlbom A, Feychting M, Koskenvuo M, et al. Electromagnetic fields and childhood cancer. *Lancet.* 1993;342:1295–1296.
10. Oak Ridge Associated Universities. *Health Effects of Low-Frequency Magnetic Fields.* Oak Ridge, Tenn: The Committee on Interagency Radiation Research and Policy Coordination; 1992. ORAU 92/F8.
11. National Radiological Protection Board. *Electromagnetic Fields and the Risk of Cancer.* Vol 3. Chilton, Didcot, UK: National Radiological Protection Board; 1992.
12. Olsen JH, Nielsen A, Schulgen G. Residence near high-voltage facilities and the risk of cancer in children. *BMJ.* 1993;307:891–895.
13. Verkasalo P, Pukkala E, Hongisto M, et al. Risk of cancer in Finnish children living close to power lines. *BMJ.* 1993;307:895–899.
14. Feychting M, Ahlbom A. Magnetic fields and cancer in children residing near Swedish high-voltage power lines. *Am J Epidemiol.* 1993; 138:467–481.
15. Ahlbom A, Feychting M. EMF and cancer. *Science.* 1993;260:13–16.
16. Washburn EP, Orza MJ, Berlin JA, et al. Residential proximity to electricity transmission and distribution equipment and risk of childhood leukemia, childhood lymphoma, and childhood nervous system tumors: systematic review, evaluation, and meta-analysis. *Cancer Causes Control.* 1994;5:299–309.
17. Miller MA, Murphy JR, Miller TI, Ruttenber AJ. Variation in cancer risk estimates for exposure to powerline frequency electromagnetic fields: a meta-analysis comparing EMF measurement methods. *Risk Anal.* 1995;15:281–287.
18. Meinert R, Michaelis J. Meta-analyses of studies on the association between electromagnetic fields and childhood cancer. *Radiat Environ Biophys.* 1996;35:15–18.
19. Myers A, Clayden A, Cartwright R, Cartwright S. Childhood cancer and overhead powerlines: a case–control study. *Br J Cancer.* 1990;62: 1008–1014.
20. Lowenthal RM, Panton JB, Baikie MJ, Lickiss JN. Exposure to high tension power lines and childhood leukaemia: a pilot study [letter]. *Med J Aust.* 1991;155:347.
21. McDowall M. Mortality of persons resident in the vicinity of electricity transmission facilities. *Br J Cancer.* 1986;53:271–279.
22. Schreiber G, Swaen G, Meijers J, Slangen J, Sturmans F. Cancer mortality and residence near electricity transmission equipment: a retrospective cohort. *Int J Epidemiol.* 1993;22:9–15.
23. Lin R, Lu P. An epidemiologic study of childhood cancer in relation to residential exposure to electromagnetic fields. In: *Project Abstracts: The Annual Review of Research on Biological Effects of Electric and Magnetic Fields. DOE/EPRI Contractor's Review Meeting, Portland, Ore, November 13–16.* Frederick, Md: Bioelectromagnetics Society; 1989:A-40.
24. Wartenberg D, Savitz DA. Evaluating exposure cutpoint bias in epidemiology studies of electric and magnetic fields. *Bioelectromagnetics.* 1993;14:237–245.
25. Hedges L, Olkin I. *Statistical Methods for Meta-Analysis.* Orlando, Fla: Academic Press; 1985.

26. Wolf F. *Meta-Analysis: Quantitative Methods for Research Synthesis*. Newbury Park, Calif: Sage Publications; 1986.
27. Fleiss J, Gross A. Meta-analysis in epidemiology, with special reference to studies of the association between exposure to environmental tobacco smoke and lung cancer: a critique. *J Clin Epidemiol*. 1991;44:127–139.
28. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7: 177–188.
29. Colditz GA, Burdick E, Mosteller F. Heterogeneity in meta-analysis of data from epidemiologic studies: a commentary. *Am J Epidemiol*. 1995;142:371–382.
30. Greenland S. Invited commentary: a critical look at some popular meta-analytic methods. *Am J Epidemiol*. 1994;140:290–296.
31. Rosenthal R. The "file drawer" problem and tolerance for null results. *Psychol Bull*. 1979; 86:638–641.
32. Cooper H. Statistically combining independent studies: a meta-analysis of sex differences in conformity. *J Pers Soc Psychol*. 1979;37:131–146.
33. Fulton J, Cobb S, Preble L, Leone L, Forman E. Electrical wiring configurations and childhood leukemia in Rhode Island. *Am J Epidemiol*. 1980;111:292–296.
34. Savitz DA, Wachtel H, Barnes FA, John EM, Tvrdik GJ. Case-control study of childhood cancer and exposure to 60-Hz magnetic fields. *Am J Epidemiol*. 1988;128:21–38.
35. London S, Thomas D, Bowman J, Sobel E, Cheng T-C, Peters J. Exposure to residential electric and magnetic fields and risk of childhood leukemia. *Am J Epidemiol*. 1991;1349: 923–937.
36. Coleman M, Bell C, Taylor H, Primic-Zakelj M. Leukaemia and residence near electricity transmission equipment: a case-control study. *Br J Cancer*. 1989;60:793–798.
37. Fajardo-Gutierrez A, Garduno-Espinosa J, Yamamoto-Kimura L, et al. Residence close to high-tension electric power lines and its association with leukemia in children. *Bol Med Hosp Infant Mex*. 1993;50:32–38.
38. Petridou E, Kassimos D, Kalmanti M, et al. Age of exposure to infections and risk of childhood leukaemia. *BMJ*. 1993;307:774.
39. Tomenius L. 50-Hz electromagnetic environment and the incidence of childhood tumors in Stockholm County. *Bioelectromagnetics*. 1986;7:191–207.
40. Wertheimer N, Leeper E. Electrical wiring configurations and childhood leukemia in Rhode Island [letter]. *Am J Epidemiol*. 1980;111: 461–462.
41. Linet MS, Hatch EE, Kleinerman RA, et al. Residential exposure to magnetic fields and acute lymphoblastic leukemia in children. *N Engl J Med*. 1997;337:1–7.
42. Michaelis J, Schuz J, Meinert R, et al. Combined risk estimates for two German population-based case-control studies on residential magnetic fields and childhood acute leukemia. *Epidemiology*. 1998;9:92–94.
43. Petridou E, Trichopoulos D, Kravaritis A, et al. Electrical power lines and childhood leukemia: a study from Greece. *Int J Cancer*. 1997;73: 345–348.
44. Tynes T, Haldorsen T. Electromagnetic fields and cancer in children residing near Norwegian high-voltage power lines. *Am J Epidemiol*. 1997;145:219–226.
45. Wartenberg D. Leukemia and exposure to magnetic fields [letter]. *N Engl J Med*. 1997; 337:1471.
46. Stevens RG. Leukemia and exposure to magnetic fields [letter]. *N Engl J Med*. 1997; 337:1472.
47. Levallois P, Gauvin D. Leukemia and exposure to magnetic fields [letter]. *N Engl J Med*. 1997; 337:1472–1473.
48. Gochfeld M. Leukemia and exposure to magnetic fields [letter]. *N Engl J Med*. 1997; 337:1473.
49. Neutra RR. Leukemia and exposure to magnetic fields [letter]. *N Engl J Med*. 1997; 337:1474.
50. Linet MS, Tarone RE. Leukemia and exposure to magnetic fields [response to letters]. *N Engl J Med*. 1997;337:1474.
51. Michaelis J, Schuz J, Meinert R, et al. Childhood leukemia and electromagnetic fields: results of a population-based case-control study in Germany. *Cancer Causes Control*. 1997;8: 167–174.
52. National Research Council. *Possible Health Effects of Exposure to Residential Electric and Magnetic Fields*. Washington, DC: National Academy Press; 1997.