NOTICE

**CERTAIN DATA
CONTAINED IN THIS
DOCUMENT MAY BE
DIFFICULT TO READ
IN MICROFICHE
PRODUCTS.**

FINAL REPORT

ASA Conference on Radiation and Health VIII

*Health Effects of Electric and Magnetic Fields:
Statistical Support for Research Strategies*

Copper Mountain Conference Center
Copper Mountain, Colorado

DISCLAIMER

July 9-13, 1989

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Chair: Michael E. Ginevan
Risk Focus Division, Versar, Inc.

The 1989 ASA Conference on Radiation and Health is partially supported by grants received from the Human Health and Assessment Division, U.S. Department of Energy and the Electric Power Research Institute.

May 1990

ASA
1989

i/ii

MASTER
American Statistical Association

1429 Duke Street, Alexandria, Virginia 22314-3402
(703) 684-1221

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

ORGANIZING COMMITTEE

Michael E. Ginevan, Chair
Risk Focus Division, Versar, Inc.

Lee A. Rosen, Vice Chair
National Institutes of Health

John C. Bailar
McGill University and U.S. Public Health Service

Peter Groer
Oak Ridge Associated Universities

Emanuel Landau
American Public Health Association

Mays Swicord
Food and Drug Administration

Anthony D. Thrall
Electric Power Research Institute

Lee L. Decker, Conference Coordinator
American Statistical Association

FELLOWS

Necip Doganaksoy
Union College

Christiana Maria Drake
University of Washington

Andrew E. Gelman
Harvard University

Michael A. Kelsh
University of California, Los Angeles

Nuala Ann Sheehan
University of Washington

SPONSORS

American Statistical Association

U. S. Department of Energy
Human Health and Assessment Division

Electric Power Research Institute

TABLE OF CONTENTS

TUTORIALS

Introduction to Extremely-Low-Frequency Electric and Magnetic Fields <i>Thomas S. Tenforde</i> , Battelle, Pacific Northwest Laboratory	1
Some Common Terminology in Statistics and Epidemiology <i>Richard P. Chiacchierini</i> , Food and Drug Administration	5

EPIDEMIOLOGICAL REVIEW I

Studies of Cancer and Residential Exposure <i>Charles Poole</i> , Epidemiology Resources, Inc.	11
Case Control Studies: Frequentist vs. Bayesian Points-of-View <i>Marvin Zelen</i> , Harvard School of Public Health	15

EPIDEMIOLOGICAL REVIEW II

Health Effects of Electric and Magnetic Fields: Occupational Studies <i>Gilles Theriault</i> , McGill University	17
Exposure Models in Retrospective Cohort Studies <i>Jack D. Sahl</i> , Southern California Edison Co. and UCLA School of Public Health <i>Michael A. Kelsh</i> , UCLA School of Public Health and Los Angeles County Natural History Museum Foundation	25

EXPERIMENTAL STUDIES I

A Summary of Health Research Related to 60 Hz Electromagnetic Fields <i>Lee A. Rosen</i> , National Institutes of Health	37
Human Laboratory Studies on the Effects of 60 Hz Fields <i>Mary R. Cook</i> , Midwest Research Institute	45

EXPERIMENTAL STUDIES II

Exposure Assessment for Power Frequency Electric and Magnetic Fields <i>T. Dan Bracken</i> , T. Dan Bracken, Inc.	49
Statistical Support for Research Strategies <i>Leeka I. Kheifets</i> , Electric Power Research Institute	53
Discussion: <i>Richard P. Chiacchierini</i> , Food and Drug Administration	59

SPECIAL PROBLEMS I

What Is Dose?	61
<i>Mays L. Swicord</i> , Food and Drug Administration	
Relative Potency Can Be Used to Evaluate ELF Health Risks	67
<i>Clay E. Easterly</i> , Oak Ridge National Laboratory	
Discussion: <i>Peter G. Groer</i> , Oak Ridge Associated Universities	73

SPECIAL PROBLEMS II

Multiple Comparisons in Statistical Analysis	75
<i>Abdelmonem A. Afifi</i> , UCLA School of Public Health	
Investigating Cellular Responses to EMF: Can Statistics Help?	83
<i>Anthony D. Thrall</i> , Electric Power Research Institute	
Discussion: <i>Richard L. Anderson</i> , Statistical Consultants, Inc.	89

PANEL ON RESEARCH STRATEGIES

Meta-Analysis	93
<i>Kay T. Kimball</i> , Baylor College of Medicine	
Research on the Health Effects of Electromagnetic Fields: Science, Uncertainty, and Stopping Rules	99
<i>John C. Bailar</i> , McGill University School of Medicine and U.S. Dept. of Health and Human Services	
Bayes Methods for Combining Dose-Response Studies	107
<i>William H. DuMouchel</i> , BBN Software Products Corporation	

RESEARCH POLICY AND RISK ASSESSMENT

Assessing, Managing, and Communicating About the Possible Risks of Exposure to 60 Hz Fields	113
<i>M. Granger Morgan</i> , Carnegie Mellon University	
Electric Fields: Discussion and Summary (Does the shape of the shovel really matter?)	119
<i>Michael E. Ginevan</i> , Risk Focus Division, Versar, Inc.	

APPENDICES

Appendix I: Conference Program	127
Appendix II: List of Participants	131

INTRODUCTION TO EXTREMELY-LOW-FREQUENCY ELECTRIC AND MAGNETIC FIELDS

T. S. Tenforde

Battelle, Pacific Northwest Laboratory
Richland, Washington

The interaction with living systems of electromagnetic fields in the extremely-low-frequency (ELF) range below 300 Hz will be summarized briefly in this paper. In materials with the electrical and magnetic properties of living tissues, these fields have a long wavelength (5000 m) and skin depth (150 m). As a consequence, in their interactions with humans and other living organisms ELF fields behave as though they are composed of independent electric and magnetic field components. This "uncoupling" of the orthogonal electric and magnetic components of an ELF field is commonly referred to as the "quasi-static approximation," which permits the radiating properties of the field to be neglected in describing its interaction with living organisms.

The electric and magnetic components of an ELF field have several distinctly different features in their interactions with humans and other living organisms (Tenforde and Kaune, 1987). First, the electrical conductivity of tissue is approximately 14 to 15 orders of magnitude greater than that of air at ELF frequencies. Consequently, the body behaves like a good electrical conductor in ELF electric fields. As a result, an electrical charge is developed on the surface of a living object in an external ELF field, but the electric field penetrates into the body only to a very limited extent ($E_{in} / 10^{-7} E_{out}$). The significant conductivity of the body relative to air also has the important consequence that the local field in air near regions of the body with small radii of curvature (e.g., the head or the fingertips) is larger by more than an order of magnitude than the external field at a distance of several meters from the body. This phenomenon is referred to as the "field enhancement effect," and it greatly complicates the dosimetry of ELF electric fields in the proximity of the body. In contrast, the magnetic component of an ELF field does not induce a surface charge and it penetrates the body without significant attenuation since the magnetic permeability of tissue is nearly identical to that of air. As a consequence, the dosimetry of ELF magnetic fields in and near a living organism is relatively simple in comparison to the dosimetry of ELF electric fields. Another feature of ELF electric and magnetic fields that differs is the pathway of induced currents within the body. Induced electric currents in tissue flow predominantly in the direction of an external ELF electric field, whereas the electric currents induced in tissue by an external ELF magnetic field circulate in a plane that is orthogonal to the direction of the applied field.

A final point to be made regarding the physical interactions of ELF fields applied to the body through air is that their interactions are primarily nonthermal in nature. This conclusion follows from the fact that the highest electric field level that can be induced in tissue by an externally applied ELF field in air is less than one V/m. Larger fields cannot be induced in tissue because the externally applied electric field would have to exceed the dielectric breakdown threshold of air. The maximum rate of joule heating of tissue by these weak induced fields is on the order of a microdegree per second, which is about five orders of magnitude less than the basal metabolic rate of humans. Because of the efficient heat conduction and convection processes in humans and other living organisms, this rate of tissue heating cannot be detected by presently available techniques.

A careful review of the published literature has demonstrated that the biological effects of ELF fields are directly correlated with the current density induced in tissue (Tenforde, 1989). A few biological effects have been observed to occur in response to induced current densities that are comparable to the naturally occurring currents that

flow in the body as a result of endogenous electrical activity of excitable tissues such as the heart and brain, i.e., 0.1 to 10 mA/m². These effects include the visual phenomena known as electrophosphenes and magnetophosphenes, and effects on the circadian rhythm in melatonin synthesis by the pineal gland. At higher levels of induced current density in the range of 10 to 100 mA/m², a variety of alterations in tissue and cellular properties have been reported to occur in response to the application of ELF fields, including beneficial effects such as the facilitation of bone fracture reunion. With induced current densities in the range of 100 to 1000 mA/m², thresholds for neuronal and neuromuscular effects are exceeded. Finally, at levels above 1 A/m² the induced currents in tissue can produce severe, and potentially fatal, respiratory and cardiac effects.

The major challenge in ELF field research at the present time is the elucidation of mechanisms by which exposure to extremely weak ELF fields can result in reproducible biological effects. A large number of physical and electrochemical models have been proposed in which the cell membrane is viewed as playing a primary role in transducing the weak signals presented by induced ELF electrical currents in tissue. A growing body of experimental evidence suggests that electrochemical events initiated at the membrane surface by circulating pericellular currents can alter ion binding to membrane macromolecules and influence ligand-receptor interactions at the cell surface (e.g., the binding of hormones, growth factors, etc.). These field interactions at the cell surface can trigger transmembrane phenomena involving alterations in ion transport and changes in the electroconformational states of membrane proteins. Events initiated at the inner membrane surface in response to these transmembrane signals can, in turn, influence the cytoplasmic concentrations of biologically important "second messengers" such as calcium ions and cyclic nucleotides that regulate macromolecular synthesis and control cellular growth and functional states. The molecular details of this cascade of transduction events that carry ELF field signals from the extracellular milieu into a living cell remain to be revealed by careful experimentation. However, the present state of knowledge strongly implicates the cell membrane as a site of ELF field transduction and signal amplification (Tenforde and Kaune, 1987).

The existence of intracellular responses to ELF field transduction is illustrated by several lines of experimental evidence. For example, it has been shown that ornithine decarboxylase levels are increased in several lines of cultured cells after exposure to a 60 Hz electric field (Byus et al., 1987). This enzyme is essential for polyamine biosynthesis, and is activated by a number of chemicals that bind to receptors at the cell surface and stimulate cell proliferation. One example is the class of tumor promoters known as phorbol esters, and the results of Byus et al. (1987) suggest that ELF field interactions may alter cellular biochemistry in a manner similar to these compounds. Phorbol esters have been shown to produce large elevations in ornithine decarboxylase activity via their binding to membrane-associated phosphokinase C receptors and the subsequent production of new messenger RNA specific for ornithine decarboxylase (Verma et al., 1986).

In other studies involving dipteran salivary gland cells, it has been demonstrated that exposure to either pulsed or sinusoidal electromagnetic fields leads to altered messenger RNA transcription patterns (Goodman et al., 1983; 1987; Goodman and Henderson, 1986). This effect is accompanied by a significant change in the spectrum of cellular proteins synthesized by the exposed cells relative to control cells (Goodman and Henderson, 1988). A total of 248 polypeptides in the control cells were resolved by two-dimensional gel electrophoresis, while 326 were observed in cells exposed to a 72 Hz pulsed magnetic field. The polypeptides synthesized in the dipteran salivary gland cells were specific to the characteristics of the ELF field to which these cells were exposed, with various polypeptides being either enhanced in quantity or suppressed relative to those observed for unexposed cells.

These recent experimental findings provide a useful set of clues regarding the biochemical events that occur in response to ELF field signals transmitted from the cell

surface into the cytoplasm. The further elucidation of these signal transduction mechanisms constitutes a major challenge for future research on ELF field interactions with living systems.

REFERENCES

- Byus, C. V., Pieper, S. E., and Adey, W. R. The effects of low-energy 60 Hz environmental electromagnetic fields upon the growth-related enzyme ornithine decarboxylase. Carcinogenesis 8:1385-1394 (1987).
- Goodman, R. and Henderson, A. S. Sine waves enhance cellular transcription. Bioelectromagnetics 7:23-29 (1986).
- Goodman, R. and Henderson, A. S. Exposure of salivary gland cells to low-frequency fields alters polypeptide synthesis. Proc. Natl. Acad. Sci. (USA) 85:3928-3932 (1988).
- Goodman, R., Abbot, J., and Henderson, A. S. Transcriptional patterns in the X chromosome of sciara coprophila following exposure to magnetic fields. Bioelectromagnetics 8:1-7 (1987).
- Goodman, R., Bassett, C. A. L., and Henderson, A. S. Pulsing electromagnetic fields induce cellular transcription. Science 220:1283-1285 (1983).
- Tenforde, T. S. and Kaune, W. T. Interaction of extremely-low-frequency electric and magnetic fields with humans. Health Phys. 53:585-606 (1987).
- Tenforde, T. S. Biological interactions and human health effects of extremely-low-frequency magnetic fields. In: Extremely-Low-Frequency Electromagnetic Fields: The Question of Cancer, pp. 291-318. Eds.: B. W. Wilson, R. G. Stevens, and L. E. Anderson. Battelle Press, Columbus, OH (1989).
- Verma, A. R., Pong, R.-C., and Ericksen, D. Involvement of protein kinase C activation in ornithine decarboxylase gene expression in primary culture of newborn mouse epidermal cells and in skin tumor promotion by 12-O-tetradecanoylphorbol-13-acetate. Cancer Res. 46:6149-6155 (1986).

SOME COMMON TERMINOLOGY IN STATISTICS AND EPIDEMIOLOGY

Richard P. Chiacchierini
Food and Drug Administration
Rockville, Maryland

ACCURACY: The closeness of an observation or statistic derived from one or more observations to the "true value."

ALTERNATIVE HYPOTHESIS: A numerical statement concerning the parameters of one or more distributions that is mutually exclusive to the null hypothesis. Sometimes called the research hypothesis because it is, in most cases, the hypothesis that the investigator would like to prove.

ANALYSIS OF VARIANCE (ANOVA): A widely-used class of methods in which the potential sources of variation of observations are estimated and compared to determine if the various levels of classifications of interest to the investigator are statistically different from one another. These methods can be used to determine equality of treatment means adjusted for cofactors, or they can be used to estimate components of variance attributable to each random source of variation.

ASSOCIATION: A numerical relationship between two or more sets of variables such that knowledge of one set yields knowledge about the other set. Such relationships may or may not be causal.

There are five factors to be considered in evaluating whether or not an association observed in an epidemiologic study is causal.

1. The strength of the association: A strong association is more likely to be causal than a weak one.
2. Specificity of the association: The exposure to the supposed causative agent in every diseased patient tends to support causality.
3. Dose-response relationship: In a causal relationship, increased exposure to the agent usually produces an increased risk in the exposed population.
4. Consistency of the association: A strong similarity between the disease distribution and the distribution of the agent in time, space, or population subclassifications tends to support causality.
5. Confirmation of findings of association: The repeated finding of the association in several studies of different populations tends to support causation.

BIAS: Mathematical definition—the difference between the expected value (mean) of a statistic or estimator and the "true value" of the population parameter being estimated. Nonmathematical definition—one or more conditions in the collection of observations that modify the observations such that estimators or statistics derived therefrom have expected values that deviate from the "true values" of the population parameters.

BIAS, OBSERVATIONAL: A condition present in the person recording the observation or in the measurement instrument that prevents an accurate recording of the data point.

BIAS, RESPONSE: A condition present in a study subject that affects the accurate recording of the response. Particularly, in epidemiologic questionnaire studies where recall of events is frequently more intense in diseased individuals than in controls, this may lead to an inaccurate evaluation of a response factor.

BIAS, SELECTION: A condition present in the assignment of subjects to study groups that allows an imbalance to occur such that too many of one type of subject are assigned to a particular group thereby producing a more (or less) favorable outcome for that group.

CASE-CONTROL STUDY: An epidemiologic investigation in which the frequency of exposure to selected risk factors of interest among a group of diseased people (cases) is compared to the frequency of exposure among one or more groups of nondiseased people (controls).

CATEGORICAL DATA: Observations that are classes of distinct groups, usually recorded as discrete numerical values; e.g., classes of dead or alive can be recorded as 0 or 1 respectively. Likewise, white, hispanic, black, native American, etc. can be assigned values 0, 1, 2, 3, etc., also known as qualitative data or discrete data.

CAUSATION: A condition in which a situation, event, or agent produces an effect in an outcome variable of study.

COHORT STUDY: An epidemiologic study in which the frequency of morbidity or mortality from a specific disease of interest in a group exposed to a suspected risk factor is compared to that in a group of unexposed people.

CONFIDENCE INTERVAL: A type of interval estimation. If we let the confidence be $(1 - \alpha)$, then over a large number of samples from a defined population from each of which a confidence interval is computed, the "true parameter" will lie in the fraction $(1 - \alpha)$ of the computed intervals. Note this does not imply for a given computed interval that there is $(1 - \alpha)$ probability that the "true parameter" lies in the interval. Once the interval is computed, the probability that the "true parameter" lies in the interval is either 0 (if it does not) or 1 (if it does).

CONFOUNDING: When two or more factors or sources of variation in a study are obtained in such a way that their independent effects on the outcome variable are sufficiently mixed that they cannot be separated, the two (or more) factors are said to be confounded.

CONTINGENCY TABLE: A cross classification of categorical data arranged in a table (of 2 or more dimensions) such that each row, column, etc. expresses the frequency of occurrence in the sample of sample subjects possessing the characteristics defined by the row, column, etc. variables; e.g., a 2x2 contingency table of disease and exposure to a certain factor may be given by:

		Factor		Total
Disease	+	a	b	a+b
	-	c	d	c+d
		a+c	b+d	n

CORRELATION: A linear relationship between two or more sets of variables. A linear association. Correlation, like association, does not imply causality.

CONTRIBUTION (PROBABILITY DISTRIBUTION): A mathematical function that assigns a probability to the occurrence of a specific value or range of specific values that can be assumed by the random variable.

DOSE RESPONSE STUDY: An investigation that attempts to statistically define the functional relationship between a response (usually a disease incidence or prevalence) and the dose of a specific agent.

EPIDEMIOLOGY: The study of the distribution and determinants of disease in human populations.

ESTIMATION: The process by which unknown parameters of a function or distribution are given approximate values (called estimates) based on available information.

FOLLOW-UP: A process in epidemiology by which study subjects are tracked and observations of variables of interest are made over time. Follow-up has two critical features: completeness and duration. Completeness refers to the proportion of the study sample followed. Duration refers to the length of time the sample is followed.

HYPOTHESIS: A numerical statement concerning one or more parameters of a probability distribution or mathematical model.

HYPOTHESIS TESTING: Statistical procedures that allow the selection between two mutually exclusive hypotheses concerning one or more parameters of a probability distribution or mathematical model.

INCIDENCE (OF A DISEASE): The total number of new cases of a given disease that come into being during a specified period of time in a given population.

INCIDENCE RATE: The total number of new cases of a given disease that come into being during a specified period of time (usually a year) in a given population divided by the average size of the population over that time.

INFERENCE: The act of making a decision or evaluation concerning one or more characteristics or properties of a population based on information obtained from a sample. Reasoning from a part to the whole.

LATENT PERIOD: The delay between the exposure to a disease causing agent and the appearance or manifestation of the disease.

MATCHING: The method by which study groups are made comparable by the deliberate or ex post facto alignment into classes of extraneous factors.

MEAN: A measure of central tendency. A population mean is obtained as the sum (or integral) of each value of the random variable weighted by its probability of occurrence (or element of probability). A sample mean is simply the sum of all observations divided by the total number of observations.

MEDIAN: That observation (sample) or value of the random variable (population) for which 50% of the observations lie below the given value and 50% of the observations lie above the given value.

MORBIDITY: Any departure from a state of physiological or psychological well-being.

NONPARAMETRIC METHODS (DISTRIBUTION-FREE METHODS): Statistical procedures that are derived in such a way that they do not rely on the form of the underlying distribution.

NORMALITY: Pertaining to the Gaussian or normal distribution, the standard form of which is given by

$$f(x) = (1/\sqrt{2\pi}) \exp(-(x-\mu)^2/2\sigma^2)$$

NULL HYPOTHESIS: A numerical statement that the results observed in a study, experiment, or test are no different from that which may have occurred by chance alone.

ODDS RATIO (CROSS-PRODUCT RATIO, RELATIVE ODDS): If a contingency table is given by

	Exposed	Not Exposed
Disease	a	b
Nondisease	c	d

then the odds-ratio, given by ad/bc , measures the association between disease and exposure. If the disease is rare, the odds-ratio in case-control studies is an estimate of the relative risk.

PARAMETER: A mathematical variable that appears in an underlying probability distribution or mathematical model. "True parameter values" are usually unknown and must be estimated from the data.

PERSON-YEAR: A unit of measurement obtained by summing the lengths of time (one year) for each person and used as a denominator in rate calculations.

POPULATION: The entire set of subjects that are the object of a study or investigation.

POWER (STATISTICAL): The probability of rejecting a null hypotheses which is false.

PRECISION: The closeness of an observation or statistic to the mean derived from repeated sampling of the same distribution (or population).

PREVALENCE: The total number of cases of a given disease or other condition in a given population at a specific time.

PREVALENCE RATE: The total number of cases of a given disease or other condition in a given population at a specific time divided by the size of the population at that time.

PROSPECTIVE STUDY: An epidemiologic study in which study subjects are followed into the future to determine their morbidity, mortality, or exposure to risk factors.

QUALITATIVE DATA: Categorical data.

QUANTITATIVE DATA: Measurement data that can assume any value on some continuous scale.

RANDOM VARIABLE: A variable whose values follow a probability distribution.

RANDOMIZATION: A process by which study subjects are assigned to experimental test conditions such that each subject has an equal chance of being selected for each condition.

REGRESSION: Statistical procedures that allow the selection of the best numerical relationship among two or more variables. Types of regression include: simple linear regression that is a straight-line fit to data points observed in a two-dimensional grid; multiple regression that is basically linear regression in n-space; polynomial regression that is a polynomial-curve fit to data points observed in a two-dimensional grid; nonlinear regression that fits functions other than polynomials to data; and

logistic regression that is a categorical data method fitting functions to data transformed by the logistic transformation.

RELATIVE RISK: The ratio of the risk of disease or death among an exposed group of study subjects to the risk among an unexposed group.

RETROSPECTIVE STUDY: An epidemiologic study in which data concerning the morbidity, mortality, or exposure to risk factors is obtained by looking back in time from the present.

SAMPLE: Any subset from a population.

SIGNIFICANCE LEVEL: The probability (usually chosen by the experimenter) that the test statistic will assume a value that will lead to a rejection of the null hypothesis when the null hypothesis is true. If the experimenter selects a probability value a priori (such as 0.05 or 0.01), then that value becomes the nominal significance level. If the experimenter reports the lowest probability consistent with the observation of the test statistic value, this is called the attained significance level.

STATISTIC: A value that is computed from sample data or any function of a random variable.

STATISTICS: A pure and applied science of creating, developing, and applying methods such that the uncertainty of inductive inferences may be evaluated.

SURVIVAL ANALYSIS: The development and application of statistical models and methods for analyzing data representing life times, waiting times, or in a general sense, occurrence times to selected events.

SYSTEMATIC ERROR: Bias.

TYPE I ERROR: The rejection of the null hypothesis when it is true.

TYPE II ERROR: The failure to reject the null hypothesis when it is false (actually, this is the rejection of the alternative hypothesis when it is true).

VARIABLE: Any attribute, phenomenon, event, or measure that can assume different values.

VARIANCE: A measure of dispersion of a set of observations defined for a sample as the sum of squares of deviations from the mean divided by one less than the total number of observations.

REFERENCES

Armitage, P. (1971) Statistical Methods in Medical Research. John Wiley and Sons, New York.

Cochran, W.G. (1963) Sampling Techniques (2nd Edition). John Wiley and Sons, New York.

Conover, W.J. (1971) Practical Nonparametric Statistics. John Wiley and Sons, New York.

Dixon, W.J. and F.J. Massey, Jr. (1969) Introduction to Statistical Analysis (3rd Edition) McGraw-Hill Book Co., New York.

Koopmans, L.H. (1981) An Introduction to Contemporary Statistics. PWS Publishers, Boston.

Katz, S. and N.L. Johnson (eds) (1988) Encyclopedia of Statistical Sciences. Vols. 1-9. John Wiley and Sons, New York.

Last, J.M. (ed.) (1983) A Dictionary of Epidemiology. Oxford University Press, New York.

Lilienfeld, A.M., E. Pedersen, and J.E. Dowd. (1967) Cancer Epidemiology: Methods of Study. The Johns Hopkins Press, Baltimore.

MacMahon, B. and T.F. Puch. (1970) Epidemiology. Little, Brown, and Company, Boston.

Snedecor, G.W. and W.G. Cochran. (1967) Statistical Methods. Ames.

Steele, R.G.D. and J.H. Torre. (1960) Principles and Procedures of Statistics. McGraw-Hill Book Company, New York.

STUDIES OF CANCER AND RESIDENTIAL EXPOSURE

Charles Poole
Epidemiology Resources Inc.
Chestnut Hill, Massachusetts

Exposure assessment presents perhaps the most important methodologic challenge to researchers investigating the association between cancer and residential exposure to electric and magnetic fields. In the first study, Wertheimer and Leeper (1979) used a coding scheme based on the size and configuration of electric power transmission and distribution lines, and on the distance of these lines from the home. Because shielding attenuates electric fields (Kaune et al. 1987, Savitz 1987), the wiring codes are best interpreted as indicative of magnetic field strength. Validation studies (Wertheimer and Leeper 1979, Kaune et al. 1987, Savitz 1987, Savitz 1988) have shown the codes to exhibit a reasonable degree of predictive value.

Three studies (Tomenius 1986, Savitz et al. 1988, Severson et al. 1988) also made magnetic field measurements in or near the home. In none of these studies were the associations with cancer stronger for the measured fields than for the wiring codes, as would be expected under the causal hypothesis if the use of field strength measurements reduced the degree of nondifferential exposure misclassification. This observation led Savitz et al. (1988) to question the premise that spot measurements are better than wiring codes as indicators of long-term exposure.

The exposure range is another important feature of each study. In the United States, where distribution lines tend to be strung overhead, studies contain higher proportions of persons in the upper exposure categories than in Europe, where distribution lines tend to be buried in a configuration that reduces magnetic field strength. Their limited exposure ranges hinder the informativeness of the European studies (Coleman et al. 1985, Myers et al. 1985, McDowall 1986, Tomenius 1986).

A final point with respect to exposure assessment is the complete absence thus far in the published literature of an attempt to incorporate exposures from all sources into a single scale. This problem is typified by the categorization of the literature into "residential studies" and "occupational studies." Exposures from electric appliances and other sources, both inside and away from the home, can be appreciable. It is difficult to say whether this source of exposure misclassification would be differential or nondifferential between the cases and controls in any given study. Presumably, it would present the potential for greater bias in studies of adults than in studies of children.

Confounding is a particularly frustrating problem in studies of childhood cancers, the etiologies of which are almost completely unknown. In the studies that examined large numbers of potential confounders (Savitz et al. 1988, Severson et al. 1988), none were found to be important. The search for a confounder to explain the associations in the residential studies has included factors, such as traffic density near the home (Roth 1985, Savitz 1988), for which the causal hypotheses with respect to cancer are at least as tenuous as the magnetic field hypothesis itself.

The analytic methods in several studies had unusual shortcomings. In Wertheimer and Leeper's first report (1979), estimates of effect were absent but computable from the data presented. In a subsequent study (Wertheimer and Leeper 1982), the authors used a statistic they called the "C-ratio," which appears to be distantly related to the exposure-odds ratio for matched data. Perhaps the most disturbing analytic procedure has been for some authors to base their analyses on counts of residences rather than people. When the number of residences per subject differs between cases and controls, residence

counts can be misleading. This problem rendered one study (Fulton et al. 1979) essentially noncontributory because the residences at birth and diagnosis were included for the cases, whereas only the birth residences of the controls were included. It would be preferable, but expensive, to collect information on all residences and to combine the data into exposure scales for examining various hypotheses with respect to induction time.

The control selection procedure in one study (Savitz et al. 1988) raised some interesting methodologic issues. The controls were selected by random-digit dialing after the close of an eight-year period of case ascertainment. Recognizing that exposure prevalence may have changed over this period of time, the investigators restricted the controls to those who were living in the study area at the time their matched cases were diagnosed. This procedure made the controls more residentially stable than the cases and, therefore, created a possibility for selection bias. Because residentially mobile controls (i.e., those who moved away from Denver) were excluded from eligibility and families in high-exposure homes appeared to be more mobile, the proportion of controls in the high-exposure category may have been biased downward. The effect of this bias would have been to inflate the associations with both wiring codes and measured fields.

The control-selection procedure in the same study (Savitz et al. 1988) had a direct influence on another potentially important source of bias: missing data, especially for measured field strengths. The authors (Savitz et al. 1988) conjectured that missing data produced a bias toward underestimation of effect. An examination of the data from this study, however, suggests little or no bias from missing field strength data.

Improved methodology could improve the state of epidemiologic research of the possible link between magnetic fields and cancer. The most important needs are to improve exposure assessment, especially by considering all sources of exposure; to widen the exposure range by targeting towns through which transmission lines run; to include larger numbers of cases of specific cancers; to select controls concurrently with case ascertainment; to collect information on all residences; and to incorporate this information into useful exposure scales that afford flexibility in examining alternative induction time hypotheses.

REFERENCES

- Coleman, M., Bell, C.M.J., Taylor, H-L. et al. Leukaemia and electromagnetic fields: A case-control study. Presented at the International Conference on Electric and Magnetic Fields in Medicine and Biology, London, December, 1985.
- Fulton, J.P., Cobb, S., Preble, L. et al. Electrical wiring configurations and childhood leukemia in Rhode Island. *Am. J. Epidemiol.* 1980;111:292-296.
- Kaune, W.T., Stevens, R.G., Callahan, N.J. et al. Residential magnetic and electric fields measured over 24-H periods. Final report to the New York State Department of Health, Power Lines Project 21082-18, 1987.
- McDowall, M.E. Mortality of persons resident in the vicinity of electricity transmission facilities. *Br J Cancer* 1986;53:271-279.
- Myers, A., Cartwright, R.A., Bonnell, J.A. et al. Overhead power lines and childhood cancer. Presented at the International Conference on Electric and Magnetic Fields in Medicine and Biology, London, December, 1985.
- Roth, H.D. An evaluation of published studies analyzing the association of carcinogenesis with exposure to magnetic fields. Final report to the Electric Power Research Institute, Project 1824-03, 1985.

Savitz, D.A. Case-control study of childhood cancer and residential exposure to electric and magnetic fields. Final report to the New York State Department of Health, Power Lines Project 218217, 1987.

Savitz, D.A. Supplement to contractor's final report: Case-control study of childhood cancer and residential exposure to electric and magnetic fields. Report to the New York State Department of Health, Power Lines Project 218217, 1988.

Savitz, D.A., Wachtel, H., Barnes, F.A. et al. Case-control study of childhood cancer and exposure to 60 Hz magnetic fields. Am. J. Epidemiol. 1988;128:21-38.

Severson, R.K., Stevens, R.G., Kaune, W.T. et al. Acute nonlymphocytic leukemia and residential exposure to power frequency magnetic fields. Am. J. Epidemiol. 1988;128:10-20.

Tomenius, L. 50 Hz electromagnetic environment and the incidence of childhood tumors in Stockholm County. Bioelectromagnetics 1986;7:191-207.

Wertheimer, N., Leeper, E. Electrical wiring configurations and childhood cancer. Am. J. Epidemiol. 1979;109:273-284.

Wertheimer, N., Leeper, E. Adult cancer related to electrical wires near the home. Am. J. Epidemiol. 1982;111:345-355.

CASE CONTROL STUDIES: FREQUENTIST VS. BAYESIAN POINTS-OF-VIEW

Marvin Zelen
Harvard School of Public Health
Boston, Massachusetts

Case Control Study methodology is an important research strategy for ascertaining if there is a statistical relationship between an adverse health event and a potential causal factor. Although the theoretical development shows that it is a valid procedure for evaluating relationships, in practice it is difficult to carry out. The theory requires that one has a random sample of cases (individuals experiencing an event) and controls. This is rarely done. As a result, many case control studies become enmeshed in controversy. On the other hand, one can adopt a Bayesian point-of-view that incorporates prior information about exposure to a causal factor. For example, most people will agree that 30-60% of U.S. adult males are smokers. Such prior information, whether it is subjective or objective, can be incorporated into the Bayesian analysis of case control studies. This talk discussed both frequentist and Bayesian approaches to case control studies. New statistical software for Bayesian applications to case control studies were illustrated.

HEALTH EFFECTS OF ELECTRIC AND MAGNETIC FIELDS: OCCUPATIONAL STUDIES

Gilles Theriault
McGill University
Montreal, Canada

Following Wertheimer 1979 report (1) of a possible association between exposure to EMF and cancer, researchers turned toward people occupationally exposed to fields of higher intensities (electrical workers) to verify the raised hypothesis.

EXPLORATORY STUDIES OF LEUKEMIA

Between 1982 and 1988, no less than 11 short papers originating from several countries and gathering information from several registries, have been published (2-12).

As seen from Figure 1, the great majority of these studies have observed an excess of leukemia among electrical workers although in many instances, this excess does not reach statistical significance. Excesses were higher for acute myeloid leukemia.

CASE-CONTROL STUDIES

In 1985-86, three case-control studies of leukemia have come up with unexpected findings.

In a group of underground coal miners, Gilman (13) reported that 25 or more years of underground mining, a surrogate of EMF exposure cause a statistically significant increased risk for all leukemia, chronic lymphocytic leukemia, and myelogenous leukemia.

Similarly, in a matched case-control study conducted in naval shipyard workers with the intent of studying the relationship between exposure to ionizing radiation and leukemia, Stern (14) observed that leukemia was associated not with ionizing radiation but with work as electricians and welders, occupations known to have high exposure to EMF.

In an attempt to confirm the putative association between exposure to gamma radiation in houses and leukemia, Flodin (15) scrutinized residential and occupational exposures of 59 alive cases of acute myeloid leukemia (AML) and 354 controls. He unexpectedly observed a highly significant risk of AML for electrical workers. These three case-control studies tend to support the existence of an excess of leukemia among electrical workers.

Three remarkable case-control studies of central nervous system cancer have been reported.

In comparing the occupational history and exposure to EMF based on panel rating, between 951 brain cancer deaths and the same number of random non-cancer controls, Lin (16) observed elevated odd ratios for electricians, electric and electronic engineers, and utility workers. He also noted the presence of a dose-response relationship.

Comparing the occupational history of 435 brain cancer deaths with that of 386 non-brain cancer deaths, Thomas (17) observed elevated odd ratios for workers exposed to microwave-radio frequency radiation, the manufacture and repair of electronic equipment (MF/RF), as well as for electric-electronic workers.

In a population based case-control study of 202 males who died from gliomas in 1969-1978, Speers (18) observed an increased risk for brain cancer among male workers

employed in occupations associated with electricity or electromagnetic fields. This increase was dose related.

COHORT STUDIES

In a search for stronger evidence of an excess of leukemia associated with occupational exposure to EMF, there have been no less than 7 cohort studies conducted so far (19-24, see Tables 1 and 2). Several occupational groups are considered. Sample size are reasonably large and observation periods extend several time over 20 years. Although only one study (23) observed a clear excess of AML, most reported risk ratios of 1.0 and over. Same observations are made for brain cancer. This not only does not rule out the possibility of an association between these two cancers and exposure to EMF, but in view of the dilution of exposure in these large cohorts studies, it tends to support it. Cohort studies have also raised the possibility of skin melanoma to be in excess among electrical workers.

ISSUES OF SPECIAL INTEREST

Welders are exposed to intense magnetic fields. In an extensive review of cancer studies in welders, Stern (25) reported that lung cancers were in excess in most studies but not leukemia. He proposed that this observation goes against the hypothesis of an association EMF—leukemia.

Aluminum production workers and chloralkali workers are exposed to strong DC currents. There have been no less than 6 comprehensive mortality studies of these workers (26-31).

Although significant excesses of acute leukemia and lympho/reticulosarcoma were observed in only one of them, it is remarkable that most SMRs are over 100.

IN SUMMARY

EMF—Leukemia

1. Pooled results from exploratory studies among electrical workers show an elevated risk for all leukemia (RR 1.18) and for AML (RR 1.46)
2. Elevated OR have been reported in three case-control studies (underground coal miners, naval shipyard workers, population based research) for all leukemia, AML as well as lymphocytic leukemia.
3. Cohort studies have yielded non constant results, some showing no excess, other showing excesses that were not statistically significant.
4. Findings among workers exposed to strong DC magnetic fields compatible with an excess of leukemia of same order of magnitude as observed for electrical workers.
5. No excess of leukemia among welders.

EMF—Brain Cancer

1. Elevated OR in association with occupations entailing exposure to EMF have been reported in 3 case-control studies.
2. As for leukemia cohort studies have yielded non constant results; some showing no excess, other (4/8) showing excesses that were not statistically significant.
3. Findings from workers exposed to strong DC magnetic fields compatible with findings from cohort studies of electrical workers.

EMF—Skin Melanoma

1. Excesses of skin melanoma were reported in 3 out of 4 cohort studies of electrical workers.
2. In a study of eye cancer in England and Wales, Swerdlow observed a remarkably high proportional registration ratio for electrical and electronics workers, which in his opinion meant an increase of eye melanoma in this group of workers.

OVERALL CAUSALITY ASSESSMENT

The consistency of the results is not obvious. Although similar observations in occupational groups have been published from different countries and from different sources of information, from one study to another, the cancer type in excess is not always the same and the occupations in which this excess is observed vary greatly from one study to another. Furthermore, most of the studies reported had positive findings and are likely to be an incomplete selection of studies in which occupational exposures have been investigated. The strength of the association is moderate to weak. Among workers, the risk can be estimated at 1.18 for all leukemia and 1.46 for acute myeloid leukemia (32). In terms of specificity of the disease, it seems that 3 cancer types have come up in more than a few studies. They are acute myeloid leukemia, brain cancer, and skin melanoma. Most of these studies are plagued with several weaknesses: sample sizes are small; several analyses are based on proportional mortality or incidence ratios; exposure to electromagnetic fields has been assessed indirectly without precise measurements and often based only on occupational titles; and confoundance has never been considered.

The biological plausibility of an association cancer versus EMF has long been the object of debate. Previous laboratory research has been negative. There has been an important breakthrough recently, particularly with regard to cell membrane windowing and ACTH mediated effects that seem to indicate that exposure to low levels EMF of 60 Hz is likely to interfere with mechanisms at the cellular as well as hormonal levels in biological tissues (33).

REFERENCES

1. Wertheimer, N., Leeper, E. Electrical wiring configurations and childhood cancer. Am. J. Epidemiol. 1979; 109:273-284.
2. Milham, S. Mortality from leukemia in workers exposed to electrical and magnetic fields. N. Engl. J. Med. 1982; 307:249.
3. Wright, W.E., Peters, J.M., Mack, T.M. Leukemia in workers exposed to electrical and magnetic fields. Lancet 1982; 2:1160-4.
4. McDowall, M.E. Leukemia mortality in electrical workers in England and Wales. Lancet 1983; 1:246.
5. Coleman, M., Bell, J., Skeet, R. Leukemia incidence in electrical workers. Lancet 1983; 1:982-3.
6. Howe, G.R., Lindsay, J.P. A follow-up study of a ten-percent sample of the Canadian Labour Force. 1. Cancer mortality in males. 1965-73. JNCI 1983; 70:37-44.
7. Morton, W., Marjanovic, D. Leukemia incidence by occupation in Portland - Vancouver Metropolitan Area. Am. J. Ind. Med. 1984; 6:185-205.
8. Pearce, N.E., Sheppard, A.R., Howard, J.K., Fraser, J., Lilley, B.M. Leukemia in electrical workers in New Zealand. Lancet 1985; 1:811-2.
9. Milham, S. Silent keys: leukemia mortality in amateur radio operators. Lancet 1985; 1:811.
10. Milham, S. Mortality in workers exposed to electromagnetic fields. Env. Health Perspect. 1985; 62:297-300.
11. Calle, E., Savitz, D.A. Leukemia in occupational groups with presumed exposure to electrical and magnetic fields. N. Engl. J. Med. 1985; 313:1476-7.
12. Linet, M., Malker, H., McLaughlin, J., Weiner, J., Stone, B.J., Blot, W., Ericsson, J., Fraumeni, J.F. Leukemias and occupation in Sweden: a registry-based analysis. Am. J. Ind. Med. 1988; 14:319-330.
13. Gilman, P.A., Ames, R.G., McCawley, A. Leukemia risk among U.S. white male coal miners. JOM 1985; 27:669-671.
14. Stern, F.B., Waxweiler, R.A., Beaumont, J.J. et al. A case-control study of leukemia at a naval nuclear shipyard. Am. J. Epidemiol. 1986; 123:980-992.

15. Flodin, U., Fredriksson, M., Axelson, O., Persson, B., Hardell, L. Background radiation, electrical work, and some other exposures associated with acute myeloid leukemia in a case-referent study. Arch. Env. Health 1986; 41(2): 77-84.
16. Lin, R.S., Dischinger, P.C., Conde, J., Farrell, K.P. Occupational exposure to electromagnetic fields and the occurrence of brain tumours: an analysis of possible associations. JOM 1985; 27:413-9.
17. Thomas, T.L., Stolley, P.D., Sternhagen, A. et al. Brain tumor mortality among men with electrical and electronics jobs: a case-control study. JNCI 1987; 79:233-238.
18. Speers, M.A., Dobbins, J.G., Miller, V.S. Occupational exposures and brain cancer mortality: a preliminary study of East Texas residents. Am. J. Ind. Med. 1988; 13:629-638.
19. Vagero, D., Olin, R. Incidence of cancer in the electronics industry: using the new Swedish Cancer Environment Registry as a screening instrument. Brit. J. Ind. Med. 1983; 40:188-92.
20. Olin, R., Vagero, D., Ahlbom, A. Mortality experience of electrical engineers. Brit. J. Ind. Med. 1985; 42:211-2.
21. Vagero, D., Ahlbom, A., Olin, R., Sahlsten, S. Cancer morbidity among workers in the telecommunications industry. Brit. J. Ind. Med. 1985; 42:191-5.
22. Tornqvist, S., Norell, S., Ahlbom, A., Knave, B. Cancer in the electric power industry. Brit. J. Ind. Med. 1986; 43:212-3.
23. Milham, S. Increased mortality in amateur radio operators due to lymphatic and hematopoietic malignancies. Am. J. Epidemiol. 1988; 127:50-54.
24. Guberman, E., Usel, M., Raymond, L., Tissot, R., Sweetman, P.M. Disability, mortality, and incidence of cancer among Geneva painters and electricians. Brit. J. Ind. Med. 1989; 46:16-23.
25. Stern, R.M. Cancer incidence among welders: possible effects of exposure to extremely low frequency electromagnetic radiation (ELF) and to welding fumes. Environ. Health Perspect. 1987; 76:221-229.
26. Milham, S. Mortality in aluminum reduction plant workers. JOM 1979; 21:475-480.
27. Andersen, A., Dahlberg, B.E., Magnus, K., Wannag, A. Risk of cancer in the norwegian aluminum industry 1982; 29:295-298.
28. Rockette, H.E., Vincent, C.A. Mortality studies of aluminium reduction plant workers: potroom and carbon departments. JOM 1983; 25:349-557.
29. Gibbs, GW. Mortality of aluminum reduction plant workers, 1950 through 1977. JOM 1985; 27:761-770.
30. Mur, J.M., Moulin, J.J., Meyer-Bisch, C., Massin, N., Coulon, J.P., Loulergue, J. Mortality of aluminium reduction plant workers in France. Int. J. Epidemiol. 1987; 16:257-264.
31. Barregard, L., Jarvholm, B., Ungethum, E. Cancer among workers exposed to strong static magnetic fields. Lancet 1985; Oct. 19:892.
32. Coleman, M., Beral, V. A review of epidemiological studies of the health effects of living near or working with electricity generation and transmission equipment. Int. J. Epidemiol. 1988; 17:1-13.
33. Aldrich, T.E., Easterley, C.E. Electromagnetic fields and public health. Env. Health Perspect. 1987; 75:159-171.

TABLE 1

Cohort studies of electrical workers (males)

ICD 9th rev	leukemia 204-207	AML 205.1	brain 191,192	skin melanoma 172,187.7	all cancers 800-999
Vagero, RR	<1.0	<1.0	<1.0	1.35* (59)	1.15* (1855)
Olin, SMR	0.9(2)	-(0)	1.0(2)	3.2(3)	0.5* (24)
Vagero, SIR	-(0)	-(0)	1.0(5)	2.5*(8)	1.03 (102)
Tornqvist, RR p. linemen	1.3(10)	<1.0(1)	1.5(13)	NR	1.1 (236)
p. station operators	1.0(16)	<1.0(3)	1.0(17)	NR	1.0 (463)
Milham, SMR a. radio operators	1.24(36)	1.76*(15)	1.39(29)	NR	0.89* (741)
Guberan, SMR	1.43(2)	-(0)	1.54(2)	0.91(1)+	1.1: (52)
SIR	1.25	-(0)	1.18(2)	0.48(1)	0.92 (78)

* statistically significant at 95%

+ all skin

NR not reported

TABLE 2

COHORT STUDIES OF ELECTRICAL WORKERS

Cancer site	ICD-code 9th rev	Vagero 1983	Olin 1985	Vagero 1985	Törnqvist 1986 Linamen station oper	Milham 1988	Gaberan 1989
All cancers	800-999	1.15* (1855)	0.5 (24)	1.03 (102)	1.1 (236)	0.89* (741)	1.14 (52)
Oropharynx	146	2.30* (7)	(0)	(0)	NR	NR	(0)
Lung and bronchus	162.2, .3, .4, .5, .8, .9	1.52* (273)	0.4 (4)	0.9 (10)	0.7 (17)	0.66* (209)	0.95 (13)
Skin melanoma	172, 187.7	1.35* (59)	3.2 (3)	2.5* (8)	NR	NR	0.91 (1)*
Skin excluding melanoma	173	< 1	(0)	1.6 (4)	1.5 (8)	NR	0.91 (1)*
Brain and other n. s.	191, 192	< 1	1.0 (2)	1.0 (5)	1.5 (13)	1.39 (29)	1.54 (2)
All lymphatic and hemato- poietic tissue	200-208	< 1				1.23* (89)	
Leukemia	238.4	< 1	0.9 (2)	(0)	1.3 (10)	1.24 (36)	1.43 (2)
Acute myeloid	204-207 205.1	< 1	(0)	(0)	NR	1.76* (15)	(0)
Lympho-reticulo sarcoma	200.0, .1	< 1	(0)	(0)	NR	0.47 (5)	(0)
Hodgkin	201	< 1	(0)	(0)	NR	1.23 (5)	2.50 (2)
Non Hodgkin lymphoma	200, 202	< 1	(0)	1.4 (4)	NR	NR	1.00 (1)
Lymphoid	202, 203	< 1	(0)	(0)	NR	1.62* (43)	(0)

* statistically significant at 95%

+ all skin

NR not reported

FIGURE 1
LEUKEMIA RISKS AMONG ELECTRICAL WORKERS

* some overlapping

EXPOSURE MODELS IN RETROSPECTIVE COHORT STUDIES

Jack D. Sahl

Southern California Edison Company
UCLA School of Public Health, Epidemiology Department
Los Angeles, California

Michael A. Kelsh

UCLA School of Public Health, Epidemiology Department
Los Angeles County Natural History Museum Foundation
Los Angeles, California

"Cheshire-Puss, she began, rather timidly, 'would you tell me, please, which way I ought to go from here?' 'That depends a good deal on where you want to get to', said the Cat. 'I don't much care where—', said Alice. 'Then it doesn't matter which way you go', said the Cat."

—Lewis Carroll, *Alice in Wonderland*, 1934

Epidemiologic research can be described as an exercise in measurement of exposure, measurement of potential confounders, and the measurement of disease outcome (Rothman, 1986). The goal of epidemiologic research then is to obtain accurate and precise measurements in order to provide valid and precise measures of effect. Developing valid and precise estimates of exposure also enhances the public health value of the work in that mitigation strategies and cost-benefit analysis can be evaluated. Finally, good exposure assessment may lead to an improved understanding of the underlying biologic mechanisms of interaction between the toxic agent and exposed people. This paper is about making decisions and designing studies to improve exposure assessment in retrospective cohort and case-control studies of occupational environments.

Epidemiologic research generally follows these steps: Defining an "exposure protocol" and an "outcome protocol," selecting an optimal study design, conducting the study, epidemiologic evaluation, statistical evaluation, and study interpretation. The focus of this paper is on the steps used in the process of assigning useful exposure-category flags to individuals in the study, sources of error in this process, and the consequences of the decisions and errors in the resulting epidemiologic measures of effect. A brief review of commonly used procedures for exposure assessment in occupational studies is also provided.

In simple terms, our problem is in deciding who is "exposed" and who is "not exposed." The goal is to develop and implement a study design that produces an exposure "flag," which represents a valid biologic model and for which we can evaluate the sources, direction, and magnitude of error. This requires that our decision-making process be made explicit, that it be based on biologic and epidemiologic principles, and that we create a "gold" standard for the exposure variable. The "gold" standard provides an image toward which we aim, an essential aspect of achieving goals.

A key part of the study design is the selection of a statistical model used in the analysis. The selected model should take into consideration the basic components of the exposure variable. In addition, however, to matching the statistical model to the biologic properties of the exposure variable, the statistical model needs to address the overall goals of the study, and the inclusion and treatment of other relevant covariates (Vandenbroucke, 1987; Robbins and Greenland, 1986; Breslow et al., 1983; Siemiatycki and Thomas, 1981). Many of the statistical issues involved in data analysis are discussed in other papers of this manuscript.

EXPOSURE PROTOCOL

Defining an exposure protocol is the first step of the exposure assessment process. This includes evaluating the biological basis of the presumed exposure-outcome relation and exposure model building and evaluation. This provides the image that you want your exposure variable to reflect. This is an iterative process that should result in a definable goal for the study exposure variable.

The development of exposure protocol is analogous to "the protocol" used in clinical trials to assign study subjects to experimental ("exposed") and placebo ("unexposed") groups. In the context of occupational health research, this type of careful classification of subjects is applied to historical data, and with random allocation assumptions, the study is viewed as if you were doing a clinical trial, randomly allocating workers into your exposed and unexposed groups (Robins, 1987).

The development of the exposure protocol should be an iterative process. You first specify the biologically relevant exposure parameter and then use available data to update your prior expectations (Leamer, 1978; Murphy, 1978). When needed, it may be appropriate to develop additional data before undertaking the planned research. This decision is framed by the goals of the study, where you would obviously elect to get additional information if the goals of the epidemiologic study could not be achieved without the improved specification of the exposure variable.

Biologic Coherence

The goal of the exposure assessment is to achieve coherence between the resultant exposure variable and a proposed biologic model, taking into consideration the relation between exposure and dose, and the form of the dose-response curve. Various dose response models may be appropriate, depending on the exposure of interest. A major part of this "biologic coherence" is the consideration of the relevant temporal aspects of exposure, including exposure start date, end date, duration, age of the individual, and the relation of these times to the onset of disease (see Figure 1) (Breslow and Day, 1987; Checkoway, Pearce, Crawford-Brown, 1989; Rothman, 1981).

Dose is usually defined in terms of the presence or levels of the agent at the biologic site of action. Estimating dose then requires knowledge of the relation between external exposure and internal dose. For power-frequency electric and magnetic fields the frequency and amplitude of the field, and the size and shape (i.e., orientation) of the individual determine the relation between external fields and internal fields or induced currents (Kaune and Forsythe, 1988; Tenforde and Kaune, 1987). Tissue specificity has been known for chemical and physical agents for many years (Gehring, Watanabe, and Blau, 1977). If this proves to be the case for power-frequency electric and magnetic fields, then the assessment of exposure process for epidemiologic studies would be vastly improved.

The temporal aspects of exposure are critical to the process of study design and study analysis. There are two components to this issue. There are biologic considerations and engineering or process considerations.

The biologic considerations for temporal aspects of exposure assessment include the relation between magnitude of exposure and time to develop the disease (induction period), differing individual susceptibility at different ages, and the time period between initiation of irreversible disease and onset of signs and symptoms, often labeled the latent period (see Figure 1) (Johnson, 1986; Rothman, 1981).

Engineering or process considerations for temporal aspects of exposure assessment involve cohort and period effects (Kupper et al., 1985; Susser, 1982). In this case

exposures or dose may be influenced by introduction of new technologies, personal protective apparatus, evolution of work practices, or administrative controls at different time periods. An individual's exposure would then be determined by when they started work, where they worked, how they worked, how long they worked, and the influence of prior exposure on subsequent exposure.

Exposure Model Building

Exposure model building refers to the iterative process by which an exposure protocol is developed for the epidemiologic study. In this case, the protocol provides a template with which to compare the exposure histories of individuals in order to make decisions regarding their exposure status. While we had previously discussed the relation between biologic processes and disease, and time of exposure and disease, here the decisions are more practical: What is a sufficient duration of exposure to call the person exposed? Should we weight early exposure time equally with more recent exposure time? What is the lag period between exposure and outcome?

Answering these and similar questions will assist the design team in developing an exposure protocol that is based on a biologic and toxicologic foundation. Answers to these questions also depend on the particular characteristics of the exposure under study and the availability of relevant information. The goal should be to develop a flexible protocol with the potential to address competing biological or toxicological models of disease occurrence, while optimizing the amount of exposure information available.

Continued refinement of the exposure protocol, until it is "good enough," will help in selecting an appropriate design, help hone the goals and objectives of the proposed study, and assist in an evaluation of measurement error during the data evaluation phase.

Once the exposure model has been adequately refined, the development of the study design is a practical process in which available resources and technical tools are matched to define the criteria used to select an exposure flag for each individual in the study. The study design determines what the exposure variable will be, how it will be created, and how it will be used in the epidemiologic and statistical evaluation. The previously finished "exposure protocol" forms the basis of these decisions.

It is important to emphasize the advantage of developing the exposure protocol as part of the design phase, and not at the end of a study in the data analysis phase. The process of developing an exposure protocol should highlight what kind of information you need, what information is available, and what compromises one will have to make in exposure assessment, data analysis, and interpretation. On a more fundamental level, the development of exposure protocol will assist in assessing whether the proposed research will have the ability to provide new information. This is useful to know at the beginning of a research project.

SURVEYS OF EXPOSURE PROTOCOLS

Various exposure protocols have been utilized in occupational health research, although often the exposure assessment is not formally defined in terms of exposure protocol, nor are the assumptions regarding exposure made explicit. The purpose of this section is to briefly review traditional exposure assignment strategies, and discuss the relevant assumptions required under these strategies.

Job Titles as Exposure Surrogates

Having decided what the definition of exposure is going to be, the crux of the problem is defining exposure in the relevant time period, usually years past. Sufficient industrial hygiene monitoring data is usually not available. Investigators tend to rely on occupational "job titles" as a proxy of quantitative data.

The occupational job title, for either individuals or groups, is often used as a surrogate estimator for exposure in epidemiologic studies. These are either self-reported on questionnaires, reported by "next-of-kin" on the death certificate, or taken from the personnel system of the employer. Once the job title is known, variations of "Expert Panels" are convened to decide which "job titles" reflect which study exposure categories (Baumgarten, Siemiatycki, and Gibbs, 1983; Bourbonnais, Meyer, and Theriault, 1988; Eskenazi and Pearson, 1988; Goldberg, Siemiatycki, and Gerin, 1986; Illis et al., 1987; Shalat, Christiani, and Baker, 1987; Stewart, Tonascia, and Matanoski, 1987; Turner, Schumacher, and West, 1987).

This task is made difficult because people tend to have many different jobs over their lifetime. Even if the researcher has access to the complete job history (which is not usually the case), there is often doubt as how best to use the information (Gamble and Spitas, 1976; Zielhuis, 1985). Some of the most important assumptions required when using job titles as exposure surrogates are listed in Table 1. These assumptions are often not met.

A refinement of the use of job titles is to create a "Job Exposure Matrix" (Hinds, Kolonel, and Lee, 1985; Hoar, 1982; Hoar, 1981; Hsieh, Walker, and Hoar, 1983; Magnani et al., 1988; Siemiatycki, Gerin, and Hubert, 1981). Here quantitative or qualitative assessments are made for each job title in regards to a series of exposures and these are arrayed in a matrix. Once again, "expert panels" are convened to decide which exposures belong to which job (see Figures 2 and 3).

When used properly the Job Exposure Matrix can reduce misclassification of exposure, usually by removing jobs truly not exposed from the "exposed job title" list. This system is more effective when the jobs being evaluated are in the same industry, or better, in the same company.

Enhancements on the Job Exposure Matrix is a Job-Task Exposure Matrix (see Figures 2 and 3). Here the individual tasks being performed are evaluated for exposure potential. Additional information on the task duration, effects of season, and changes in exposure potential that occur over time can also be included in these systems.

MEASUREMENT ERROR

One part of epidemiologic and statistical analysis of occupational health data involves the assessment of how error in the exposure variable influences the epidemiologic measures of effect. Measurement error in the exposure variable results from procedural error and construct error. Procedural error refers to sloppy measurement techniques, limited precision and reliability of the instrument, and the use of surrogate parameters that do not reflect the exposure of interest. Construct error refers to the use of an inappropriate biologic model.

Measurement error usually exists in epidemiologic data and is characterized as a misclassification problem. The task is to determine the magnitude and the direction of the error. Here the previously developed exposure protocol is a valuable tool. The magnitude of the influence will be a function of the degree of the sum of errors. The direction of the influence, or misclassification, will depend on the relation of the amount of error relative to the outcome status of the individual (Barron, 1977; Copeland et al., 1977; Flegal, Brownie, and Haas, 1986; Greenland and Kleinbaum, 1983; Greenland, 1982; Kromhout et al., 1987; Lagakos, 1988).

SUMMARY

The overall process of exposure assessment is complex. While we usually cannot achieve the goal of perfect "exposure classification," it is important to define what the

goal is. This helps us to improve our study design, and in the process of evaluation, to better estimate the magnitude and direction of measurement errors on the epidemiologic measures of effect.

REFERENCES

- Barron, B.A. 1977. The effects of misclassification on the estimation of relative risk. Biometrics 33:414-418.
- Baumgarten, M., J. Siemiatycki, and G.W. Gibbs. 1983. Validity of work histories obtained by interview for epidemiologic purposes. Am. J. Epid. 118:583-591.
- Bourbonnais, R., F. Meyer, and G. Theriault. 1988. Validity of self reported work history. Brit. J. Ind. Med. 45:29-32.
- Breslow, N.E. and N.E. Day. 1987. Modelling the relationship between risk, dose, and time. In: Statistical Methods In Cancer Research. Volume 2 - The Design And Analysis Of Cohort Studies. pp. 232-270. IARC Sci. Publ. 82. IARC, Lyon (France).
- Breslow, N.E., et al. 1983. Multiplicative models and cohort analysis. J. Am. Stat. Assoc. 78:1-12.
- Checkoway, H., N. Pearce, and D.J. Crawford-Brown, 1989. Research Methods In Occupational Epidemiology. Oxford University Press, New York.
- Copeland, K.T., et al. 1977. Bias due to misclassification in the estimation of relative risk. Am. J. Epid. 105:488-495.
- Eskenazi, B. and K. Pearson. 1988. Validation of a self-administered questionnaire for assessing occupational and environmental exposures of pregnant women. Am. J. Epid. 128:1117-1129.
- Fidler, A.T., E.L. Baker, and R.E. Letz. 1987. Estimation of long term exposure to mixed solvents from questionnaire data: a tool for epidemiological investigations. Brit. J. Ind. Med. 44:133-141.
- Flegal, K.M., C. Brownie, and J.D. Haas. 1986. The effects of exposure misclassification on estimates of relative risk. Am. J. Epid. 123:736-751.
- Gamble, J. and R. Spirtas. 1976. Job classification and utilization of complete work histories in occupational epidemiology. J. Occup. Med. 18:399-404.
- Gehring, P.J., P.G. Watanabe, and G.E. Blau. 1977. Pharmacokinetic studies in evaluation of the toxicological and environmental hazard of chemicals.
- Goldberg, M.S., J. Siemiatycki, and M. Gerin. 1986. Inter-rater agreement in assessing occupational exposure in a case-control study. Brit. J. Ind. Med. 43:667-676.
- Greenland, S. and J.M. Robins. 1986. Identifiability, exchangeability, and epidemiological confounding. International J. Epid. 15:413-419.
- Greenland, S. and D.G. Kleinbaum. 1983. Correcting for misclassification in two-way tables and matched-pair studies. International J. Epid. 12:93-97.
- Greenland, S. 1982. The effect of misclassification in matched-pair case-control studies. Am. J. Epid. 116:402-406.

- Hilaski, H.J. 1985. Where occupational health reporting systems go wrong. Am. J. Ind. Med. 8:435-439.
- Hinds, M.W., L.N. Kolonel, and J. Lee. 1985. Application of a job-exposure matrix to a case-control study of lung cancer. JNCI 75:193-197.
- Hoar, S.K. 1982. Job-exposure matrices in occupational epidemiology. JNCI 69:1419-1420.
- Hoar, S.K. 1981. Epidemiology and occupational classification systems. In: Banbury Report 9. Quantification of Occupational Cancer. Peto, R. and M. Schneiderman. Cold Springs Harbor Laboratory.
- Hsieh, C.C., A.M. Walker, and S.K. Hoar. 1983. Grouping occupations according to carcinogenic potential: occupation clusters from an exposure linkage system. Am. J. Epid. 117:575-589.
- Illis, W.R., et al. 1987. Summary measures of occupational history: a comparison of latest occupation and industry with usual occupation and industry. AJPH 77:1532-1534.
- Johnson, E.S. 1986. Duration of exposure as a surrogate for dose in the examination of dose response relations. Brit. J. Ind. Med. 43:427-429.
- Kaune, W.T. and W.C. Forsythe. 1988. Current densities induced in swine and rat models by power-frequency electric fields. Bioelectromagnetics 9:1-24.
- Kromhout, H., et al. 1987. Agreement between qualitative exposure estimates and quantitative exposure measurements. Am. J. Ind. Med. 12:551-562.
- Kupper, L.L., et al. 1985. Statistical age-period-cohort analysis: a review and critique. J. Chron. Dis. 38:811-830.
- Lagakos, S.W. 1988. Effects of mismodelling and mismeasuring explanatory variables on tests of their association with a response variable. Stat. Med. 7:257-274.
- Leamer, E. 1978. Specification Searches. John Wiley & Sons, New York.
- Magnani, C., et al. 1988. Application of a job-exposure matrix to national mortality statistics for lung cancer. Brit. J. Ind. Med. 45:70-72.
- Murphy, E.A. 1978. Epidemiologic strategies and genetic factors. Int. J. Epid. 7:7-14.
- Robins, J. 1987. A graphical approach to identification and estimation of causal parameters in mortality studies with sustained exposure periods. J. Chronic. Dis. 40 Suppl. 2 139S-161S.
- Robins, J.M. and S. Greenland. 1986. The role of model selection in causal inference from nonexperimental data. Am. J. Epid. 123:392-402.
- Rothman, K.J. 1986. Modern Epidemiology. Little Brown & Co., Boston, MA.
- Rothman, K.J. 1981. Induction and latent periods. Am. J. Epid. 114:253-259.
- Shalat, S.L., D.C. Christiani, and E.L. Baker. 1987. Accuracy of work history obtained from a spouse. Scand. J. Work Environ. Health 13:67-69.
- Siemiatycki, J., M. Gerin, and J. Hubert. 1981. Exposure-based case control approach to

discovering occupational carcinogens: preliminary findings. In: Banbury Report 9. Quantification of Occupational Cancer. Peto, R. and M. Schneiderman. Cold Springs Harbor Laboratory.

Siemiatycki, J. and D.C. Thomas. 1981. Biological models and statistical interactions: an example from multistage carcinogenesis. International J. Epid. 10:383-387.

Smith, T.J. 1987. Exposure assessment for occupational epidemiology. Am. J. Ind. Med. 12:249-268.

Stewart, W.F., J.A. Tonascia, and G.M. Matanoski. 1987. The validity of questionnaire-reported work history in live respondents. J. Occup. Med. 29:795-800.

Susser, M. 1982. Period effects, generation effects and age effects in peptic ulcer mortality. J. Chron. Dis. 35:29-40.

Tenforde, T.S. and W.T. Kaune, 1987. Interaction of extremely low frequency electric and magnetic fields with humans. Health Physics 53:585-606.

Turner, D.W., M.C. Schumacher, and D.W. West. 1987. Comparison of occupational interview data to death certificate data in Utah. Am. J. Ind. Med. 12:145-151.

Vanderbroucke, J.P. 1987. Should we abandon statistical modelling altogether? Am. J. Epid. 126:10-13.

Zielhuis, R. 1985. Total exposure and workers' health. Ann. Occup. Hyg. 29:463-475.

Table 1.

Job Titles as Exposure Surrogates

-Assumptions-

- Title reflects exposure of interest
 - Fits biologic model
- No individual effects
 - Work practices
 - Engineering controls
 - Administrative controls
- Time period trends
- Individual effects
 - Age
 - Sex

Figure 1
Time-Related Factors in Exposure

(Checkoway et al., 1989)

Figure 2
Job/Task Exposure Matrix

Figure 3

Job/Task Exposure Matrix

2000-01-01
2000-01-01
2000-01-01

A SUMMARY OF HEALTH RESEARCH RELATED TO 60 HZ ELECTROMAGNETIC FIELDS

Lee A. Rosen
National Institutes of Health
Washington, D.C.

INTRODUCTION

The study of electromagnetic field health effects was originally motivated by concerns for occupational workers and residents living adjacent to high voltage transmissions lines. Research efforts have taken two tracks: epidemiological studies of residents and occupational workers, and laboratory studies ranging from the subcellular to the human volunteer.

This abstract will focus on laboratory studies of cellular, tissue, and animal responses to exposure to electric and/or magnetic fields. The following information summarizes the results that have been reported over the last twelve to fifteen years. Other presentations during this conference will address the epidemiology and human volunteer research.

Throughout this report, references will be made to electric and magnetic fields. From a physical standpoint, these fields are different, can be separately identified and measured, and can be studied separately in the laboratory. However, each is capable of inducing a current density in the organism or the cell. Induced current in turn results in the presence of electric and magnetic fields. There are no data as yet to indicate whether an organism responds differently to electric or magnetic fields. Efforts are currently under way to address the issue and its importance.

LITERATURE REVIEW

A general review of the literature of laboratory studies indicates that most systems of the intact animal are unaffected by exposure to EMF. There are data, however, that clearly demonstrate organisms are responsive to EMF exposure. For example, rats exhibit a startle response to the onset of a high strength electric field (Rosenberg et al., 1983), and reasonably accurate thresholds of electric field detection have been identified for rats (Stern et al., 1983), baboons (Orr et al., 1986) and humans (Graham and Cohen, 1985). In addition, there are indications that the nervous system is the primary site of sensitivity to EMF, summarized as follows:

- Nerve-muscle preparations from animals exposed to electric fields respond differently than preparations from sham exposed animals (Jaffe et al., 1981).
- Magnetic Fields alter the release of neurotransmitters into the synaptic space of neuromuscular junctions (Gunderson et al., 1986).
- Electrical recordings from hippocampal slices change in latency and amplitude with application of current (Bawin et al., 1986).
- Neurotransmitter concentrations in brain regions of rodents (Vasquez et al., 1988) and in the cerebral spinal fluid of primates (Walpow et al., 1987) change with exposure to electric or combined electric and magnetic fields respectively. The neurotransmitter concentration changes in rodents exposed to electric fields indicates a four hour shift in concentration peaks.
- Nighttime levels of pineal melatonin are suppressed (Wilson et al., 1981) or the time of the peak is shifted (Reiter et al., 1988) in rats exposed to electric fields. Stevens (1987) has suggested that loss of nighttime melatonin is one possible explanation for increased cancer incidence.

- Changes in daily or circadian rhythms occur in response to exposure to electric fields (rodents, Groh et al., 1987), magnetic fields (humans), and combined fields (primates, Sulzman et al., 1986).
- Baboon behavioral data indicate that after an initial startle response, most operant and social behavioral activities are not affected by extended exposure to electric fields (Rogers et al., 1986). There is a suggestion that combined EMF may affect timing in rats (Thomas et al., 1986), and that rats exposed during fetal development may express altered behavior three months following cessation of exposure (Salzinger 1987).

There are studies where a number of parameters have been tested with no overall effects of exposure to EMF, but a single or a small number of individual measurements are reported responsive to the fields. These studies, which are inconclusive and controversial, include:

- Endocrinology—Studies with rodents (Free et al., 1981) and human volunteers (Graham and Cohen, 1985) have failed to demonstrate effects of exposure to electric or combined electric and magnetic fields on growth, stress, and reproductive hormones. Serum testosterone levels were reduced in rodents (Free et al., 1981), but associated serum protein and reproductive indicators were not altered. In one laboratory corticosterone levels were reported to be altered with extended exposure (Seto et al., 1984). Results from a second laboratory suggest that changes in hormone concentrations, such as corticosterone, were related to the time of day the measurements were taken (Free et al., 1981).
- Immunology—Tests of humoral and cellular parameters and of general leukocyte metabolism in two rodent species have been negative (Morris 1987). Reports of a cytotoxic T-lymphocyte experiment indicated that exposure to electric fields alter the killing capability of the exposed cell and that there is a dose response (Lyle et al., 1988).
- Tumor development—Studies have assessed the effects of electric or magnetic fields on growth and development of tumors. Most studies, both those with well characterized facilities (Thomson et al., 1988) and those with no specific information on facility engineering features, have reported no change in tumor growth (Le Bars et al., 1983). A recent study reported no difference in tumor development; however, there was an increase in the number of tumors per tumor-bearing animals (Leung et al., 1988).
- Teratology—Results from a study with swine exposed to electric fields suggest an increased number of teratologic events in exposed animals (Sikov et al., 1987). However, a series of rodent studies conducted at two laboratories and at several different levels of field intensity failed to find similar results (Rommereim et al., 1987; Seto et al., 1984).

Finally, there is a large body of literature that reports that exposure to electric fields has no impact on the general growth and development of animals (Smith et al., 1981), on cellular genetics (Cohen et al., 1986) general pathology (Ragan et al., 1979) or reproductive activity of animals (Carsten et al., 1987). Some additional screening efforts are required for comparative data for exposure of animals to magnetic and combined fields.

Defining the basis for much of the controversy in the literature should be a major focus of this program. As shown in Figure 1, corticosterone levels measured in rats following different exposure conditions, using similar methods of measurement, show some similarities of data for the controls but significant differences in the magnitude and direction of change for exposed animals compared to controls. In all cases, field exposures were well in excess of animal perception, but there was a possibility of inhomogeneous fields in one study where the animals could actually have been shielded from exposure. Exposure paradigms also differed in the longer exposure studies, with daily periods of fields being off in one but time out only occurring two or three times weekly in the other.

Overall duration of exposure could also be a factor. Finally, evaluation of the statistical results are a factor, the conclusions of the published investigators being provided. Thus, differences could relate to experimental design, definination of dose, engineering considerations, or evaluation of statistical analysis.

Figure 1
Comparison of Corticosterone Measurements Between Laboratories

	Seto et al., 1984 (80 kV/m) Generation			Free et al., 1991 (100 kV/m) Days of Exposure			Michaelson, 1985 (100 kV/m) Exposure	
	First	Second	Third	30 D	120 D	120 D	0.5 Hr	1 Hr.
All values-ug/dl								
Sham Exposure								
Ave Value	3.58	3.35	6.40	8.28	9.08	6.44	12.0	7.51
SE	0.5	0.21	0.57	2.21	1.59	0.52	4.6	2.7
N	42	43	42	18	19	24		
Exposed								
Ave Value	5.21	4.91	6.67	4.34	5.15	7.65	32.7	13.4
SE	0.77	0.64	0.69	1.06	0.99	0.90	10.9	3.50
N	40	43	43	17	20	25		
p=	0.05	0.059	0.75	0.12	0.02	0.49	N/A	N/A
Investigator								
Conclusion:	Significant increase			No signif. change			No signif. change	

The concept of a mechanism of interaction between the applied field and biological tissue is the foremost question in this program. Physical theory predicts that field intensities associated with transmission lines are insufficient to overcome or affect the inherent cellular transmembrane potential, and are too low to induce heating. However, data are available, in addition to the animals data above, that demonstrate changes in cellular enzymes (Byus et al, 1987), lymphocyte activity (Lyle et al., 1988), ion uptake (Liburdy and Sheeler, 1988), gap junction activity (Fletcher et al., 1988), and transcription and translation activity (Goodman and Henderson, 1988) at field intensities too low to produce depolarization or heating. Various hypotheses have been presented to explain such effects including changes in membrane channel size or configuration, in the structure or action of membrane surface receptors or alterations in the structure of the membrane lipid bilayer, or possibly in the amplification of externally received signals across the membrane (Adey, 1987; Adey 1988). It has also been suggested that the cell membrane is not merely the sum of the parts, that presently available physical theories are not appropriate, and that new predictive theories may have to be developed (Rosen and Bond, 1989).

Several scientific committees have reviewed the literature, including the epidemiology data, and agree that exposure to electric and/or magnetic fields cause biological changes in living organisms. However, while there are no identifiable health hazards from acute exposure, there is insufficient information to come to any conclusion on the possible consequences of long term exposures (Rosen et al., 1988). More recent literature reports suggest that pathologies such as cancer, are possible through such

mechanisms as endocrine manipulations, shifts in circadian rhythms, or enhancement of the initiation/promotion process.

At present, the explanation and significance of the changes are only speculative. Replication is needed by independent laboratories and characterization is required to define the various effects and their significance with respect to human exposure. Without definition of a mechanism of interaction, it will be difficult to define the potential risk of exposure. Such definitions will need to include whether one component of the applied field is more biologically important than another, such as the magnetic field over the induced current, or even whether there is more than one mechanism, such as skin.

Statistical procedures are an obvious part of this research, as in any other program. Some of the unique features of the study of EMF health effects that could have a statistical basis include:

- A problem of reproducibility or robustness of reported effects;
- Experimental numbers are small because special exposure facilities are required;
- The number of parameters tested in screening studies is large, resulting in possible false positive results; and
- Many of the hypotheses being tested, for example, hormonal control of the progression stage of cancer development and dual stress considerations, require multiple, progressive samples from the experimental subject.

At this point in the program, two possibilities can be suggested, that reports of effects of EMF are actually artifacts of the experimental protocol, or statistical support to many programs is not adequate to compensate for the various weaknesses inherent in the program. It is possible that new statistical means of analysis or application of advanced forms of statistical developments, may be required, which is the point of this conference. Such support would be valuable in assisting in: (1) defining appropriate protocols for more focused studies based on specific hypotheses, (2) using available engineering and epidemiological data for defining possible definitions of does other than high, continuous exposure, including separation of electric from magnetic fields, considerations as intermittency, time above possible thresholds of activity, all of which can be validated in the laboratory before providing better direction for future epidemiological studies, (3) providing valid experimental designs of interrelated *in vitro* and *in vivo* studies of possible mechanisms of interaction to ensure that all issues of artifact and dependent variables are properly addressed, and (4) defining better, larger animal studies for experimental definition of major issues such as cancer promotion or progression, multiple step, dependent changes in the neuroendocrine and/or immunological system, or separation of the multiple components of the electromagnetic field.

In summary, more communication is required between the statistical and health science community associated with the study of potential electromagnetic field health effects.

REFERENCES

Adey, W.R., 1986, The sequence and energetics of cell membrane transductive coupling to intracellular enzyme systems. Bioelectrochem. Bioenergetics 15:447-456.

Adey, W.R., 1988, The cellular microenvironment and signaling through cell membranes. In: Electromagnetic Fields and Neurobehavioral Function. M.E. O'Connor and R.H. Lovely, eds. Progress in Clinical and Biological Research, Vol. 257, 1988. Liss, NY. pp 81-106.

Bawin, S.M., A.R. Sheppard, M.D. Mahoney, M. Abu-Assal, and W.R. Adey, 1986. Comparison between the effects of extracellular direct and sinusoidal currents on

excitability in hippocampal slices. Brain Research 362:350-354.

Byus, C.V., S.E. Pieper, and W.R. Adey, 1987, The effects of low-energy 60 Hz environmental electromagnetic fields upon the growth-related enzyme ornithine decarboxylase, Carcinogenesis, 8:253-260.

Carsten, A.L., R.D. Benz, J.W. Baum, and A.V. Kuehne, 1987, Mutagenicity and toxicity of 60-Hz magnetic and electric fields, submitted to the New York State Power Lines Project, New York State Department of Health, Albany, NY.

Cohen, M., A. Kunska, J.A. Astemborski, D. McCulloch, and D.A. Paskewitz, 1986, Effect of low-level 60 Hz electromagnetic fields on human lymphoid cells: I. Mitotic rate and chromosome breakage in human peripheral lymphocytes, Bioelectromagnetics, 7:425-434.

Fletcher, W.H., W.W. Shiu, T.A. Ishida, and W.R. Adey, 1987, A modulated microwave field and tumor promoters may inhibit cell to cell communication and cause an increased insensitivity to cytotoxic lymphokines and tumor necrosis factor, International Symposium—Biological Effects From Electric and Magnetic Fields, Air Ions, and Ion Currents Associated with High Voltage Transmission Lines, 2-5 November, Kansas City, MO (Abstract).

Free, M.J., W.T. Kaune, R.D. Phillips, and H.-C. Cheng, 1981, Endocrinological effects of high strength electric fields on rats, Bioelectromagnetics, 2:105-122.

Goodman, R. and A.S. Henderson, 1988, Exposure of salivary gland cells to low-frequency electromagnetic fields alters polypeptide synthesis, Proc. Natl. Acad. Sci. USA., 85:3928-3932.

Graham, C., and H.D. Cohen, 1985. Influence of 60 Hz Fields on Human Behavior, Physiology and Biochemistry, Final Report, submitted to the New York State Power Lines Project, New York State Department of Health, Albany, NY, 102 pages.

Groh, K., 1987, The actions of high strength 60 Hz electric fields on circadian rhythms in small rodents, International Symposium—Biological Effects From Electric and Magnetic Fields, Air Ions, and Ion Currents Associated with High Voltage Transmission Lines, 2-5 November, Kansas City, MO, Abstract.

Gunderson, R., B. Greenbaum, and M. Schaller, 1986, Intracellular recordings during magnetic field application to monitor neurotransmitter release events: Methods and preliminary results. Bioelectromagnetics, 7:271-281.

Jaffe, R.A., B.L. Laszewski, and D.B. Carr, 1981. Chronic exposure to a 60 Hz electric field: Effects on neuromuscular function in the rat, Bioelectromagnetics, 2:227-239.

Le Bars, H., G. Andre, F. Pupin and C. Labie, 1983, Research studies on biological effects of electric field; Effects on the rat, the mouse, the guinea-pig [French], Recueil de Medecine Veterinaire, 159(10):823-837.

Leung, F., D.N. Rommereim, R.G. Stevens, B.W. Wilson, R.L. Buschbom, and L.E. Anderson, 1988, Effects of electric fields on rat mammary tumor development induced by 7,12-Dimethylbenz(a)anthracene, Bioelectromagnetics Society Tenth Annual Meeting, Stamford, CT, 19-23 June, 1988, page 2 (Abstract).

Liburdy, R.P. and C.Q. Sheeler, 1988, Calcium metabolism in the lymphocyte during interaction with magnetic fields, Contractors Review - Biological Effects From Electric and Magnetic Fields, Air Ions, and Ion Currents Associated with High Voltage Transmission Lines, 30 October-3 November, Phoenix, AZ (Abstract).

Lyle, D.B., R.D. Ayotte, A.R. Sheppard, and W.R. Adey, 1988, Suppression of T-lymphocyte cytotoxicity following exposure to 60 Hz sinusoidal electric fields, Bioelectromagnetics, 9:303-313.

Michaelson, S., 1989, Final Report: Endocrine balance in Rats Exposed to 60 Hz Electric Fields. Submitted to the US Department of Energy, Washington, DC.

Morris, J.E., 1987, Immunological studies and 60 Hz electric fields: A summary, International Symposium—Biological Effects From Electric and Magnetic Fields, Air Ions, and Ion Currents Associated with High Voltage Transmission Lines, 2-5 November, Kansas City, MO (Abstract).

Orr, J.L., J.H. Lucas, H.D. Smith, and W.R. Rogers, 1986. Determination of threshold intensity for detection of 60 Hz electric fields, Bioelectromagnetics Society Eighth Annual Meeting, 1-5 June, Madison, WI, page 37 (Abstract).

Ragan, H.A., M.J. Pipes, W.T. Kaune, and R.D. Phillips, 1979, Clinical evaluations in rats and mice chronically exposed to 60-Hz electric fields, In: Biological Effects of Extremely Low Frequency Electromagnetic Fields, R.D. Phillips, M.F. Gillis, W.T. Kaune and D.D. Mahlum, (eds) 18th Annual Hanford Life Sciences Symposium, October 16-18, 1978, Richland, WA. CONF-781016, NTIS, Springfield, VA.

Reiter, R.I., L.E. Anderson, and B.W. Wilson, 1988, Effects of 60 Hz on pineal function of rats in utero, Life Science, 42:2203-2206.

Rogers, W.R., J.H. Lucas, H.D. Smith, and J.L. Orr, 1986. 60 Hz electric fields do not serve as a negative reinforcer in nonhuman primates. Bioelectromagnetics Society Eighth Annual Meeting, page 38 (Abstract).

Rommereim, D.N., W.T. Kaune, R.L. Bushbaum, R.D. Phillips, and M.R. Sikov, 1987b, Reproduction and development in rats chronically exposed to 60 Hz electric fields. Bioelectromagnetics, 8:243-258.

Rosen, L.A. and J.D. Bond, 1989, Workshop Final Report: The Role of Cooperative Phenomena in Understanding Electromagnetic Field Effects in Biological Membranes, US Department of Energy, Electric Power Research Institute and National Institutes of Standards and Technology, 15 Pages.

Rosen, L.A., J. Sahl, and J. Lee, 1988, Forum: Task force report on power frequency electric and magnetic field effects, IEEE Power Engineering Review, 8(12):3-8.

Rosenberg, R.S., P.H. Duffy, G.A. Sacher, and C.F. Ehret, 1983. Relationship between field strength and arousal response in mice exposed to 60 Hz high voltage electric fields, Bioelectromagnetics, 4:181-191.

Salzinger, K., 1987, Behavioral Effects of ELF, Final Report Submitted to the New York State Power Lines Project, New York State Department of Health, Albany, NY.

Seto, Y.J., W.P. Dunlap, F.T. Fox, S.T. Hsieh, J.R. Lyman grover, C.F. Walker, and D. Majeau-Chavgois, 1984, Central nervous system and endocrinology studies for power frequency electric field effects. Final Report, submitted to Electric Power Research Institute, Palo Alto, CA.

Sikov, M.R., D.N. Rommereim, J.L. Beamer, R.L. Buschbom, W.T. Kaune, R.D. Phillips, and L.E. Anderson, 1987. Developmental studies of Hanford miniature swine exposed to 60 Hz electric fields, Bioelectromagnetics, 8:229-242.

Smith, L.G., R.L. Bushbaum, and M.R. Sikov, 1981, Effects of continuous electric-field

exposure on four generations of mice. In: Biological Effects of High-Strength Electric Fields on Small Laboratory Animals, Interim Report, by R. Phillips, L. Anderson, and W. Kaune, DOE/RL/ 01830-T7.

Stern, S., V.G. Laties, C. Stancampiano, C. Cox, and J.O. de Lorge, 1983, Behavioral detection of 60 Hz electric fields by rats, Bioelectromagnetics, 4:215-247.

Stevens, R., 1987, Electric power use and breast cancer: A hypothesis, Amer. J. Epidemiology, 125:556-161.

Sulzman, F.M. and D.E. Murrish, 1986, Effects of Electromagnetic Fields on Primate Circadian Rhythms, Final Report submitted to the New York State Power Lines Project, New York State Department of Health, Albany, NY.

Thomas, J.R., J. Schrot, and A.R. Liboff, 1986, Low-intensity magnetic fields alter operant behavior in rats. Bioelectromagnetics, 7:349-357.

Thomson, R.A.E., S.M. Michaelson, and Q.A. Nguyen, 1988, Influence of 60 Hz magnetic fields on leukemia, Bioelectromagnetics, 9:149-158.

Vasquez, B.J., L.E. Anderson, L. Lowery, and W.R. Adey, 1988, Diurnal patterns in brain biogenic amines of rats exposed to 60 Hz electric fields, Bioelectromagnetics, 9:229-236.

Wilson, B.W., L.E. Anderson, D.I. Hilton, and R.D. Phillips, 1981, Chronic exposure to 60 Hz electric fields: Effects on pineal function in the rat, Bioelectromagnetics, 2:371-380.

Wolpaw J.R., R.F. Seegal, R.I. Dowman, and S. Satya-Murti, 1987, Chronic Effects of 60 Hz Electric and Magnetic Fields on Primate Central Nervous System Function, submitted to the New York State Power Lines Project, New York State Department of Health, Albany, NY.

HUMAN LABORATORY STUDIES ON THE EFFECTS OF 60 HZ FIELDS¹

Mary R. Cook
Midwest Research Institute
Kansas City, Missouri

Human research on 60-Hz fields presents special design and statistical problems because the effects are subtle, small, and difficult to replicate. This presentation focused on some of the problems and issues our laboratory had to address in studying human exposure to powerline frequency electric and magnetic fields.

Our basic research strategy has been to first identify which human functions appear to be affected by field exposure; second, to see if these effects can be replicated; third, to understand the specific circumstances under which they occur; and finally, to identify the underlying mechanisms.

To date, we have evaluated exposure effects in six controlled laboratory studies involving over 200 healthy young men. Exposure conditions and the general characteristics of the facility were described by Cohen, Graham, and Cook (1987).

In each study, exposure and environmental conditions were checked and verified by outside experts, the test sessions were performed under double blind conditions, and if the study included multiple exposure sessions, the sessions were counterbalanced. The findings from these studies indicated that exposure to 9 kV/m, 200 mG combined electric and magnetic fields results in a small but significant slowing of heart rate, changes in components of the electrocardiogram, and alterations in evoked potential measures of brain physiology. These effects appeared more rapidly, and were seen more clearly, when exposure to the fields was intermittent rather than constant.

Several statistical and design issues were common to all of the studies. These issues were discussed and illustrated with examples of our approach to the problem.²

Independent groups vs. repeated measures: For each study, it was necessary to decide whether to use an independent groups or a repeated measures, subject-as-own-control, design. Was one design more appropriate than the other, and if so, under what circumstances? The most important advantage of using each subject as his own control is that individual differences per se will not confound the results. However, there are also disadvantages to such a design. The potential exists for carry-over effects; using a subject in multiple sessions increases the ubiquitous problem of practice and habituation effects; subjects may show "first" and "last" day effects, (e.g., anxiety about the procedures on the first day, relief that it is finally over on the last day). Finally, using each subject as his own control brings up the issue of the best way of counterbalancing the study.

Our first exposure study (Fotopoulos et al., 1987; Graham et al., 1987) provided an excellent context in which to discuss the various trade-offs involved. In that study, multiple measures of human performance, physiology, biochemistry and subjective state were obtained on 12 male volunteers over a two-week baseline period and during four exposure test sessions. In the discussion of these issues, examples were used to illustrate how we statistically evaluated the impact of each design on the hypotheses tested in the studies.

Experimentwise error rate: Human research is very expensive, and it would be foolish not to obtain as much data as possible from each experimental session. Doing so, however, can produce a great many dependent variables. When coupled with the use of a repeated measures design, this approach can also create special problems in controlling the

experimentwise error rate. In our initial exposure study, we really did not worry about the experimentwise error rate. We felt that in a screening study it was more dangerous to say that an effect did not exist if in fact it did, than to say that an effect existed when it did not. We knew that any apparent effects would have to be confirmed in a subsequent replication study. Since the initial study, we have reduced Type 1 error by limiting analyses to the testing of specific hypotheses, by performing preliminary analyses to eliminate as many factors as possible from the design, and by grouping variables together empirically for multivariate analyses.

Protection against Type 2 error: Most of the variables we examined appeared to be unaffected by exposure to 60-Hz fields. What evidence is required to determine that no effect exists? Discussion centered on confidence as a function of the power of the test used, which depends on the number of subjects involved, the size of the effect, and the variability of the data (Cohen, 1977).

Distinguishing real from spurious effects: It is obvious that statistical significance alone is not sufficient to determine that an effect has occurred. To deal with this problem, we have used internal replications, and have carefully examined the pattern of statistically significant results for consistency with regard to the human functions being examined. Examples were given of the effectiveness of this approach.

Specific testing of assumptions: Every research design is based on a set of implicit assumptions, which may or may not be correct. Under what circumstances should these assumptions be explicitly tested? For example, when we began our work, it was generally assumed that if effects were found they would be a positive function of the intensity and duration of exposure. As our research progressed, however, we began to doubt these apparently reasonable assumptions. Exposure effects in our studies occurred more frequently at certain "windows" of electric and magnetic field stimulation, and they appeared to be related more to switching the fields on and off than to either the duration or intensity of exposure. Discussion centered on how one decides to change the direction of a research program, and how this can be accomplished within the time and cost constraints of funded research.

Control groups: The final issue concerned the fact that almost all people in our society are exposed to electric and magnetic fields. What is the appropriate research strategy to use under such circumstances? What is an appropriate control group? How do we know whether a response is or is not beyond the "normal" range, when normal ranges have been set on people who are chronically exposed to electric and magnetic fields? We have no answers to these questions.

Only a few years ago, most experts believed that exposure to 60 Hz electric and magnetic fields was unlikely to have replicable biological effects. Since then, mounting evidence in the areas of cell biology, neurochemistry, and epidemiology has led to a re-evaluation of the potential effects of powerline frequency fields. The research community is now faced with the task of determining whether, in fact, replicable biological effects exist, and if so, whether these effects constitute any risk to the public. Accomplishing these tasks requires sensitive and cost-effective experimental designs and analytical techniques to help overcome the ubiquitous problems presented by the issues discussed here.

FOOTNOTES

1. The work discussed here was funded by the Department of Energy (DE-FC01-84-CE-76246), the New York State Powerlines Project (21082-03), and the U.S. Environmental Protection Agency (CR-812805-01-0). The contents of this publication do not necessarily reflect the views of the funding agencies.

2. In this presentation, examples from the various studies were used to illustrate each problem and our approach to its solution. Due to space limitations, the points raised are summarized here. A more complete presentation is available from the author (Mary R. Cook, Ph.D., Biobehavioral Sciences Section, Midwest Research Institute, 425 Volker Boulevard, Kansas City, MO 64110.

REFERENCES

Cohen, H. D., C. Graham, and M. R. Cook (1987), "A 60 Hz Human Exposure System: Design and Dosimetry," in Hanford Life Sciences Symposium Series, No.23, Interaction of Biological Systems with Static and ELF Electric and Magnetic Fields, Pacific Northwest Laboratory, Richland, WA.

Cohen, J. (1977), Statistical Power Analysis for the Behavioral Sciences, Academic Press, New York.

Fotopoulos, S. S., C. Graham, M. R. Cook, E. Koontz, M. M. Gerkovich, and H. D. Cohen (1987), "60-Hz Field Effects on Human Neuroregulatory, Immunologic, Hematologic and Target Organ Activity," in Hanford Life Sciences Symposium Series, No.23, Interaction of Biological Systems with Static and ELF Electric and Magnetic Fields, Pacific Northwest Laboratory, Richland, WA.

Graham, C., H. D. Cohen, M. R. Cook, J. W. Phelps, M. M. Gerkovich, and S. S. Fotopoulos (1987), "A Double-Blind Evaluation of 60-Hz Field Effects on Human Performance, Physiology and Subjective State," in Hanford Life Sciences Symposium Series, No. 23, Interaction of Biological Systems with Static and ELF Electric and Magnetic Fields, Pacific Northwest Laboratory, Richland, WA.

EXPOSURE ASSESSMENT FOR POWER FREQUENCY ELECTRIC AND MAGNETIC FIELDS

T. Dan Bracken
T. Dan Bracken, Inc.
Portland, Oregon

Exposure assessment is the quantification of a subject's encounters with an environmental agent: in this case, electric and magnetic fields associated with the electric power system. The construct within which this quantification of exposures to these fields will be presented is the exposure assessment matrix. The exposure assessment matrix consists of rows for each time or age category of a subject and columns for each environment a subject experiences. The individual cells within the matrix contain information about the field levels and time spent in that environment for the particular age. Exposure within a certain time frame, say, childhood, or within a specific environment, say, occupation, can be estimated by aggregating over appropriate cells.

Several methodologies have been used to assign electric and magnetic field levels to a particular (age x environment) matrix cell. For example, residential fields and exposures have been estimated from point-in-time area measurements [1,2], 24-hour measurements [1], personal exposure measurements [3], and using surrogates such as wire codes [1,2,4,5] and empirical models [1]. Occupational exposures have been associated with point-in-time measurements in the workplace [6], personal exposure measurements [3], job classification or occupational title [7], and even by industry [7]. No matter which method is used, there are several common issues related to uncertainties in the exposure levels.

Paramount among these issues is the lack of a meaningful exposure metric. Because a clear mechanism of interaction between fields and humans that could be associated with health risk has not been identified, it is not possible to specify a parameter for characterizing dose. Consequently, the choice of field parameters for exposure assessment is somewhat arbitrary: electric field, magnetic field, or both? 60 Hz fields, harmonics, or both? peak field, average field, or field variation? For purposes of this presentation, the magnitudes of the electric and magnetic fields at 60 Hz and the lower harmonics are used to illustrate the questions that arise in the exposure assessment process for power frequency fields.

More traditional problems that contribute to uncertainty and are amenable to input from statisticians are also present. These include:

- extrapolation from present data to past conditions;
- sampling strategies;
- description of extremely skewed distributions;
- testing of surrogates;
- quantification of uncertainty in aggregate models;
- the significance of misclassification errors; and
- determination of significant differences in exposure.

Ideally in an exposure assessment a random sample can be drawn from the population in question. In the case of electric and magnetic field exposure, as well as with other environmental agents, this is generally not possible and models of exposure must suffice. The exposure model construction and validation process should rely on reliable measurements based on sound scientific and statistical principles. Considerations include instrument selection, measurement errors, sample size, duration of measurements, and elimination of possible biases.

Besides the familiar questions of measurement error and the adequacy of the sample, there is also the question of extrapolating present measurements to historical levels. In the case of childhood residential exposures where the time between purported exposure and measurements may only be a few years, the relationship between present measurements and actual exposures may be straightforward. However, for retrospective studies with cases more than a decade ago and for lifetime occupational exposures, the hypothesis that present measurements are representative of past conditions requires investigation and justification: Has the electrical distribution system changed in the intervening period? What is the evidence for long term stability of the sources associated with the fields in question? Have work practices remained the same? Has new equipment been introduced?

Measurements of both electric and magnetic fields have indicated that the frequency distributions of these fields are highly skewed. For example, the vast majority of occupational exposures, even among an "exposed" group, occur at low levels (<10 mG), while relatively little time is spent in high field levels (>100 mG) [3,8]. In these cases the time-averaged mean exposures are dominated by very few measurements and are not a good indicator of the nature of the exposures. Extremely skewed distributions have also been observed for electric field exposures [3,9]. The presence of highly skewed distributions coupled with the aforementioned lack of an exposure metric, gives rise to several questions: Do we need to characterize the rare extreme exposure values? If so, at what level, 1%, 0.1%, 0.01%? What are appropriate sampling strategies to capture the desired level of information? How can we describe these highly skewed distributions with a practical number of descriptors and retain the basic information needed to examine exposure scenarios? For example, is the arithmetic or geometric mean and its associated deviation appropriate or should we rely on the median with quantiles to give a useful description of central tendency and variation?

Another method of assigning field levels to subjects is through the use of surrogates. For residential exposures, the characteristics of the nearby distribution wiring as embedded in a wire code has been the predominant surrogate for magnetic field exposures. In occupational studies, job classification or occupational title has served as a surrogate for electric and magnetic field exposures. Much current research is directed at establishing the validity and limitations of these surrogates: Magnetic field measurements are being made in numerous residences coincident with wire coding and other observations [10,11]; magnetic field exposures are being evaluated as a function of occupational title [12]. As with using measurements to establish exposures, the use of measurements to validate surrogates gives rise to questions about measurement error, sample size, duration of measurements, etc.

Although emphasis has been placed on establishing field levels in various exposure matrix cells, knowledge of time spent in the environment associated with each cell is also of importance. This information can be collected from individuals with contemporaneous activity logbooks or from subject interviews to obtain past history. Historical allocation of time for groups can be estimated using time-budget databases or with expert panels. Because electric and magnetic field exposures can be dominated by a few sources, extensive interviews regarding proximity and use of appliances, electrical equipment and electrical facilities may be necessary.

The problems of associating data acquired in the present with past events that occur in establishing historical exposure levels also occur when establishing historical time/location information. The uncertainty in determining past exposure times or locations may be even greater than that associated with the field levels. For this reason and because the tendency has been to concentrate on field levels, the importance of assessing the validity and uncertainty of time/locations estimates cannot be over-estimated in the exposure assessment process. In fact, analysis of this element in an exposure matrix cell can help determine the precision necessary in the measurements or estimates of the physical parameters in that cell.

Exposure scenarios for identified subpopulations are generated by aggregating the time/location land field levels information over the appropriate exposure matrix cells. A typical approach is to sum the time-weighted exposure over all relevant environments and ages to produce a single value for time integrated exposure. Another possibility, given the distribution of fields within each cell, is to estimate the time when exposure exceeded certain threshold values. The lack of an identified exposure metric, of course, precludes the selection of a single methodology for aggregating exposures. However, in generating various aggregated exposures common steps occur: the subset of exposure matrix cells particular to a subpopulation is identified; a preliminary exposure assessment utilizes knowledge of field levels and time/location data to determine which of these cells will be included in the aggregation process; the exposure metric is selected and the aggregation is performed, and; the uncertainty of the aggregated exposure estimate is estimated from the uncertainties in the field and time/location data.

Having generated aggregate exposures for individuals or groups, the actual identification of exposed and unexposed subjects can be completed. In the context of epidemiological studies, this is the purpose of the whole exposure assessment process. However, even at this point, there are still questions and problems of a statistical nature. These include: What constitutes a statistically significant difference in estimated exposure? What are impacts of misclassifications of, say, residential field levels? What is the sensitivity of the analysis to relative changes in assigned field levels, to the omitted cells, to the selected exposure metric?

The exposure matrix for a laboratory study is straightforward: field level and exposure schedules are carefully controlled (e.g., 100 kV/m for 20 hours per day for 100 days). In such studies, it is the selection and interpretation of the exposures that is difficult, rather than the quantification. What level of exposure for a rat or other species scales to realistic or worst-case exposures for humans? Or even more difficult, how do you interpret in-vitro exposures in terms of human exposures? The issue of scaling between species becomes very tentative given the lack of a clear mechanism for significant effects.

The exposure assessment process for electric and magnetic fields is replete with questions and problems of a statistical nature. The purpose of this presentation is to highlight some of these and to stimulate an interdisciplinary solution that involves epidemiologists, physical scientists, engineers, and statisticians.

REFERENCES

1. Kaune, W.T., Stevens, R.G., Callahan, N.J., Severson, R.K., Thomas, D.B. Residential magnetic and electric fields. Bioelectromagnetics 1987; 8:315-35.
2. Barnes, F., Wachtel, H., Savitz, D., Fuller, F. Use of Wiring Configuration and Wiring Codes for Estimating Externally Generated Electric and Magnetic Fields. Bioelectromagnetics 1989; 19:13-21.
3. Deadman, J.E., Camus, M., Armstrong, B.G., Heroux, P., Dyr, D., Plante, M., Theriault, G. Occupational and residential 60 Hz electromagnetic fields and high-frequency electric transients: exposure assessment using a new dosimeter. American Industrial Hygiene Association Journal 1988; 48:409-19.
4. Wertheimer, N., Leeper, E. Electric wiring configurations and childhood cancer. American Journal of Epidemiology 1979; 109:273-84.
5. Wertheimer, N., Leeper, E. Adult cancers related to electrical wires near the home. International Journal of Epidemiology 1982; 11:345-55.
6. Bowman, J.D., Garabrant, D.H., Sobel, E., Peters, J.M. Exposures to extremely low frequency (ELF) electromagnetic fields in occupations with elevated leukemia rates. Applied Industrial Hygiene 1988; 6:189-94.

7. Coleman, M., Beral, V. A review of epidemiological studies of the health effects of living near or working with electricity generation and transmission equipment. International Journal of Epidemiology 1988; 17:1-13.
8. Bracken, T.D. Unpublished data from EPRI EMDEX Project (RP 2966-1), 1989.
9. Bracken, T.D. Analysis of BPA Occupational Electric Field Exposure Data. Final Report Contract DEAC 79-83 BP 36303; Bonneville Power Administration, Vancouver, WA; Jan. 1986.
10. Peters, J.M., Bowman, J.D., Thomas, D.C. Childhood Leukemia and Residential Exposure to Electric and Magnetic Fields. Proceedings of the EPRI Utility Seminar: Epidemiologic Studies of Electromagnetic Field Exposure; Minneapolis, MN; Sept. 30 - Oct. 2, 1987.
11. Mitsche, J.V. Overview of EPRI Research Program. EPRI Electrical Systems Division. Magnetic Fields Characteristics, Sources and Prediction. Proceedings of the EPRI Utility Seminar: Power Frequency Electric and Magnetic Field Exposure Assessment; Colorado Springs, CO; Oct. 12-14, 1988.
12. Bowman, J.D., Garabrant, D.H., Assessment of occupational exposures to electric and magnetic fields. Proceedings of the EPRI Utility Seminar: Epidemiologic Studies of Electromagnetic Field Exposure; Minneapolis, MN; Sept. 30 - Oct. 2, 1987.

STATISTICAL SUPPORT FOR RESEARCH STRATEGIES

Leeka Kheifets
Electric Power Research Institute
Palo Alto, California

INTRODUCTION

Residential studies of electric and magnetic fields (EMF) and cancer have employed various definitions and categorizations of exposure. They include distance to power lines, wiring configurations, power consumption, spot measurements of electric and magnetic fields, and long and frequent measurements (as frequent as once per second for one to three days).

The observed associations between cancer and EMF exposure are inconsistent across, and at times within, studies. Overall, the observed association is weak and is based on small differences (on the order of few milligauss) in average exposure categories. The preponderance of residential exposures are low with high exposures occurring rarely. (Note that exposures encountered in some occupational environments are orders of magnitude higher than residential exposures, but here too high exposures are rare). For example, Savitz (1988) found a majority of residences (90%) with average exposure values below 1 mg, with few of homes above 3 mg. Although some high peak exposures have been recorded in some residences, average exposures in houses rarely reach 10 mg.

Contrary to what one would expect, observed associations between cancer and EMF are stronger with crude surrogates of exposure than with actual measurements (Myers 1985, Tomenius 1986, Savitz 1988). Several explanations of this phenomenon have been advanced. They range from the belief that the associations reported were spurious to an assertion that wire codes are better and more stable indicators of past exposures.

The various representations of exposure show moderate to poor correlation with one another. A cross classification of two representations of exposure (wire codes and spot measurements) are presented in Figure 1. Homes with high wire codes (VHCC and OHCC combined) are rare and represent less than 30% of the residences in all studies. Note that Zaffanella (1988) specifically selected some of the homes to be in the high exposure category. Similarly high classification based on actual measurements is also rare. The misclassification pattern that emerges from this cross classification is telling. Most of the measurements in homes of low wire code are also low, but the correspondence between high wire codes and actual measurements is very poor. Based on this misclassification pattern, one could argue that wire codes are not a better predictor of long-term exposure. Because if they were, one would expect consistently high measurements in the high wire code configuration homes and more misclassification in low wire code homes due to the contribution of other sources to short-term measurements.

Sources of magnetic fields in residential environments are many and varied. Appliances generate fairly high fields but these fields are usually intermittent and fall off rapidly with distance. Transmission lines are an important source of magnetic fields in homes that are located in close proximity to them, but this occurs infrequently. Imbalances in distribution lines and ground currents have been identified as important contributors to the levels of magnetic fields present in many homes.

MEASUREMENT STRATEGIES

As an individual moves through a residential or occupational environment, the magnetic fields to which the individual is exposed vary greatly. This makes exposure assessment extremely difficult. To capture this variability, investigators have favored

increasing the duration of measurement to 24 hours or longer while maintaining sampling rates as high as once per second. This produces an enormous amount of data for each location. But some preliminary data indicates that shorter measurements (10 min. to half an hour) capture almost the same amount of variability and provide similar ranking (Figure 2 from Peters 1989, Juutilainen 1989). Therefore, a more efficient design might incorporate repeated measurements, perhaps for just a sub-sample, in place of the effort needed to retrieve the instruments left for 24 hours.

Profiles of magnetic fields in homes vary: in some homes magnetic fields are fairly stable; others exhibit high variability with no apparent pattern; still others have a distinct diurnal pattern. A possible sampling scheme is to vary the length of measurements in the house according to the pattern of variability observed during a short-term measurement.

Since an appropriate metric of exposure is not known, many types of measurements are performed. The situation is further complicated by the fact that for most individuals we will be interested in both current and previous residencies and in exposures received while away from home. Variability due to individual activities is superimposed on a background of diurnal and seasonal variability. Although diurnal patterns of magnetic fields have been observed in some homes, they are not present in others. Seasonal variations seem likely, but they have not been quantified.

Exposure could be assessed more efficiently if the hypothesized mechanism of cancer promotion were better defined, and if we already had information about exposure profiles. It might then be possible to reduce the number of measurement locations, and to measure at an appropriate sampling rate for an appropriate period of time. In addition, estimates of within- versus between-person variability are urgently needed to intelligently design future studies.

Despite its desirability, development of a standardized measurement protocol is premature. Measurements are becoming more precise and complex, with personal dosimetry planned for the next generation of studies. Nevertheless, we are still lacking the basic information needed to develop a historic exposure profile for children and adults.

The question is then, how should the measurements be allocated between different homes and other environments; should one attempt repeated measurements during different seasons, or one longer measurement; which rooms should be included within a residence and how well should each be characterized; which parameters of the fields need to be measured and under what power conditions? Better sampling strategies need to be implemented to capture complexity and variability of the magnetic fields with reasonable effort and cost.

CHOICE OF EXPOSURE METRIC AND SUMMARIZATION OF EXPOSURE

Results of recent epidemiologic studies indicate that magnetic fields may influence biologic phenomena more than electric fields. However, levels of magnetic and electric fields are correlated, making the distinction difficult. Clues from basic research have been few and difficult to interpret. To test these theories we will need to collect new data, e.g., on harmonics, geomagnetic fields, and the directionality of fields. The questions are then should we be measuring electric or magnetic fields; are we interested in a resultant or a certain axis; which frequencies should be measured?

Several statistical parameters could be used to describe this complex and variable exposure. Cumulative exposure, represented by a time-weighted average or a median, has been used most often. It has also been suggested that peak, rather than cumulative exposures, are most important. For example, we could characterize peak exposures by computing the length of time the field's magnitude is above some threshold. The choice of this threshold is crucial for this type of summary measure. Alternatively, a maximum exposure level could be of interest. Moreover, the variability of exposure, including both

frequency and the rate of change should be considered.

For each of the measurements discussed above, the choice of summary statistic is unclear. Both arithmetic and geometric means have been used. Trimmed means have not been used but should be explored. Spectral analysis might conveniently summarize this type of data and provide additional insight into the pattern of exposure.

Among the issues requiring further study are the statistical framework to aid in the choice of the estimator, the estimation of its distribution, and the implications of the interdependence of various measurements.

MODELING OF EXPOSURE

Another issue is the choice between personal dosimetry versus exposure modeling based on area measurements and time location/activity.

Personal dosimetry could capture the variability of today's exposure, but it is unlikely that personal dosimetry today could capture an exposure pattern of the past, especially for children. We have yet to determine the practicality of personal monitoring and whether the monitoring changes behavior and thus exposures. Moreover, this extrapolation back in time may be more accurate for cases than for controls.

Modeling based on area measurements and time-location activity is also fraught with problems. While area measurements of magnetic fields could be made with high precision, allocation of time to different locations is much cruder.

How can we choose between these two extrapolation scenarios: area measurements and activity patterns, which would provide more information for extrapolation into the past, versus personal dosimetry that better captures today's exposures?

Other variables can be used in models to predict magnetic field exposure. Such variables are useful when measurements are not available, as would be the case for subjects who do not participate in the study. Due to differences in wiring, grounding, and heating patterns, however, such exposure models are not easily transferable to other studies or locations. Thus they are often built upon and applied to the same data. This practice may lead to overestimating the reliability of the models.

In addition to statistical models that predict exposure, physical models are being developed. Contrasting and unifying these models will be a challenge.

ANIMAL STUDIES

Experimenters face similar difficulties in the choice of exposure parameters. Here the issue is the choice of field parameters to vary in the experiment, rather than the characterization of existing fields. Among the field parameters of interest are magnitude, harmonic content, and direction. Field polarization is suggested to best mimic human exposure.

Unlike chemical exposures, extrapolation of human exposure to EMF to animals requires scaling up. Stray fields from the highly exposed animals produce sizeable fields among the 'unexposed' animals, thus necessitating cage controls.

Finally, the number of exposure groups and the choice of exposure levels is being debated. On the one hand a standard toxicologic approach calls for the highest possible exposure to establish the existence of an effect. Only after an effect is thus established, the argument goes, should an attempt be made to estimate dose-response. On the other hand, other scientists propose the possible existence of frequency or intensity "windows" in which adverse effects occur, and thus favor testing at many exposure levels. A third

approach is to test only at levels appropriately scaled to reproduce human exposure.

Animal experiments could attempt to replicate epidemiologic findings under more controlled experimental conditions. Alternatively, they could attempt to explore the association between EMF and cancer, by examining variety of exposure levels and characteristics. Either approach is difficult. A multitude of experimental parameters, suggestions of windows of frequency and intensity, and a large number of tests performed are only a few of the statistical issues in experimental EMF research.

CONCLUSION

A large number of epidemiologic, exposure assessment, and animal studies are being designed. Statistical input in this area of research will be very timely and important.

The statistical challenge is to develop a rigorous strategy to elucidate a weak, but potentially important, association between an ill-defined outcome and a vaguely identified exposure.

Statistical input is crucial in all aspects of investigation. In particular, design of the study should include a clear definition of the problem and implementation of the most efficient sampling strategy; measurement protocols should be evaluated for their validity and reliability; in modeling of exposure we should explore uncertainty in all of the parameters and in their joint effect, as well as the predictive power and goodness of fit of the models; finally, during analysis the statistician is crucial in the definition of exposure, testing of assumptions, distinction between confirmatory and exploratory hypotheses testing, and issues of multiple testing.

REFERENCES

- Fulton, J.P., et al. 1980. Electrical wiring configurations and childhood leukemia in Rhode Island. Am. J. Epidemiol. 111:292-296.
- Juutilainen, J., et al. 1989. Measurements of 50 Hz Magnetic Fields in Finnish Homes. Research Report, Imatran Voima Oy.
- Peters, J. 1989. Private communication.
- Savitz, D.A., et al. 1988. Case-control study of childhood cancer and exposure to 60 Hz magnetic fields. Am. J. Epidemiol. 128:21-38.
- Tomenius, L. 1986. 50 Hz electromagnetic environment and the incidence of childhood tumors in Stockholm County. Bioelectromagnetics 7:191-207.
- Wertheimer, N. and E. Leeper 1979. Electrical wiring configurations and childhood cancer. Am. J. Epidemiol. 109:273-/284.
- Zaffanella, L.E. 1988. Power Frequency Magnetic Fields in Residences: A Pilot Study. Draft Report, EPRI Research/Project 2942.

Figure 1. Wire Code / Measurement Discrepancies

Rank ordering of residences by magnetic fields measured for different durations

figure 2

DISCUSSION

Experimental Studies II

Richard P. Chiacchierini
Food and Drug Administration
Rockville, Maryland

This has been an interesting couple of days. Because I do not have an intimate professional role to play in this risk assessment, my perspective is devoid of the emotion that frequently characterizes discussions and clouds the differentiation of fact from fiction. Oftentimes, such discussions generate far more heat than light. I would like to think that my view is objective, but none of us is truly objective. So, within the framework of my own biases, here is how I view the evidence.

There may be a weak association between some disease entities and some ill-defined exposure variables, which may and probably do differ between studies, observed inconsistently in several human populations. There does not appear to be compelling experimental evidence to support the association, nor is there a plausible physical or biological mechanism that can be relied upon.

Our two papers in this session do a good job of discussing major problems with exposure assessment. Dr. Bracken has indicated several areas where the statistician might assist in this process. Dr. Kheifets has provided a number of questions in this area, all of which point to a major problem that was echoed by both speakers.

"There is no exposure--assessment metric!"

While Dr. Bracken indicated a number of legitimate areas of statistical support, the statistician is being asked to assist in the estimation of a unnamed variable, the metric, whose underlying distribution is also unknown. Please excuse my naivete, but I left my crystal ball at home. The statistician can give advice in the selection process of such a metric, but until a variable is chosen for study, the actual application of statistical methodology must be put on hold. The application of statistics to the exposure--assessment problem is not yet at square one.

There are several questions that I would recommend be answered in an attempt to address the exposure--assessment problem.

1. What question do we want to answer?

What metric or combination of metrics best characterize the exposure of living whole animals to electric and magnetic fields?

So, the first activity must be the selection of highly plausible metrics that could be considered as viable exposure--assessment candidates. This might be accomplished by some high-level committee composed of representatives from physics, engineering, biology, epidemiology, statistics, and perhaps other fields. The number of metric candidates should be culled to a precious few, perhaps to no more than three or four. This will require discipline to avoid including each member's pet metric.

An example of one type of metric can be obtained from ionizing radiation. In uranium mining situations, where there is a complexity of exposure conditions, the working-level-month was developed to describe the exposure.

2. What information do we need?

The metrics chosen in answer to question 1 need to be characterized statistically by measurement under a variety of realistic electric and magnetic field conditions. If the instrumentation does not exist to make appropriate measurements, it needs to be developed. When instrumentation is available, laboratory measurements need to be developed under carefully controlled conditions. Relevant statistical estimators for the metrics need to be developed and tested in the laboratory. These must be tailored to the specific metric: continuous chronic exposure, pulsed chronic exposure, acute peak exposures, chronic peak exposures, etc. At this point a total characterization of each metric under ideal laboratory conditions must be attained. This includes a characterization of the probability density function of the metric with an identification and estimation of relevant population parameters. Any relationships between the chosen metrics should also be identified.

3. What is the relevance to real life?

A very strong attempt must be made to use the metrics to characterize the surrogate exposure variables in human studies. When this is completed to the degree possible, repeat the analyses using each metric as the exposure variable. Consistent associations should be noted.

4. It is impossible to go back 20 years and measure a metric that did not exist at that time. Stochastic models can be developed in conjunction with sound physical and engineering principles on the historical changes and their impact on the metric. Where possible, facilities or homes from the relevant era should be sampled and the metrics measured. These measures are necessary to validate the model. This modeling effort could turn out to be substantial and expensive.

These questions are probably not new, but I feel the study of exposure assessment must begin with some concrete answers to identify the relevant metric. Even if the choices in answer to question 1 turn out to be wrong, the information gained will eventually lead to the appropriate measure. There is an important role for the statistician, i.e., to collaborate with the subject matter specialists, but there must be some direction given to the effort. Once this is done, a whole arena of interactions becomes available.

WHAT IS DOSE?

Mays L. Swicord
Food and Drug Administration
Washington, D.C.

The question of what is dose is raised in the context of quantifying the relationship between exposure to extremely low frequency (ELF) electromagnetic fields and biological effects induced by such an exposure. Such a discussion necessitates a general framework of dosimetric concepts. The approach used in ionizing radiation protection and developed by the International Commission on Radiation Protection (ICRP Report 26, 1977) will be used for comparison. The objective is to identify physical quantities that characterize ELF fields and that can be related in a quantitative manner to biological effects induced by exposure.

All the physical quantities that can be used to characterize unperturbed ELF fields are radiometric quantities (external electric field strength, external magnetic field strength, or magnetic flux density) supplemented by time-related parameters, such as frequency, modulation, duration of exposure and its distribution over time (continuous, intermittent, single, or repeated).

This discussion will begin with the classical definition derived from ionizing radiation, move to the application of these concepts for hazard assessment for microwave radiation, and then focus on the various parameters of exposure to extremely low frequency electric and magnetic fields. The objective will be to show that the normal use of the term dose is not applicable to the Extremely Low Frequency (ELF) range.

Early in the use of ionizing radiation for clinical treatment the developers realized the need for an unit for inter-comparison of radiation dose. This unit, the roentgen was the unit of exposure and not absorbed dose. The limitations of this approach were soon realized and the concept of absorbed dose (the quotient of the mean energy imparted per unit mass) was introduced (ICRP Report 30, Quantitative Concepts in Radiobiology).

The concept of dose used in microwave and radiofrequency radiation protection follows the tradition of defining an absorbed quantity, which when an excess amount is absorbed, will result in harm. Quoting from the American National Standards Institute (ANSI) C95.1-1982 standard (the same text is used in the 1990 draft standard). "Dosimetry is the fundamental process of measuring physical quantities of energy or substances that are imparted to an absorbing body." In 1972, the National Council on Radiation Protection and Measurements (NCRP) convened Scientific Committee 39 to deliberate and recommend dosimetric quantities and units applicable to deliberate exposure, and recommend dosimetric quantities and units applicable to radiofrequency fields. In keeping with the NCRP recommendations the ANSI subcommittee adopted the unit-mass, time-averaged rate of energy absorption as specified in SI units of watts per kilogram (W/kg). The quantity expressed by these units is termed the specific absorption rate (SAR). Formally defined, the specific absorption rate is "the time rate at which radiofrequency electromagnetic energy is imparted to an element of mass of a biological body." The integral energy absorbed over time is the specific absorption (SA) or dose.

With this concept of dose, one merely needs to determine the amount of energy absorbed at a particular frequency by a specified tissue to assess the potential for harm. In the radiofrequency range this is not a trivial task. The wavelength in the radiofrequency range is of the same dimensional order as the human body or parts of the human body. This results in a geometrical interaction between the radiation and the absorbing body resulting in considerable variation in the energy absorbed depending on the

frequency and the geometry. Engineers have developed sophisticated modeling techniques

to calculate the absorbed energy in an object under a given exposure condition. Thus any given exposure can be related to a whole body SAR or SA (dose).

At frequencies above about 100 MHz the exposure is generally in a radiating field where the electric and magnetic fields are coupled and have a constant relationship. Here the exposure can be expressed as power density. At lower frequencies the relationship between the electric and magnetic field may vary with location. Thus it is necessary to express the exposure in terms of both the electric and magnetic field. The magnetic field has the ability to penetrate the body and induce electric fields. Thus in some situations the magnetic field may be of greater concern. In either case, electric fields or induced electric fields from magnetic fields cause charge motion (current flow) in the body resulting in the absorption of energy. In addition, the induced currents will affect cell membrane function through electrochemical processes.

Voluntary standard setting bodies have utilized the engineering calculations to establish guidelines for safe exposure to electromagnetic fields. To do so, one reviews the data base of biological effects to determine a specific absorption rate that is considered safe. Due to the lack of frequency specific biological effects data this analysis is done without regard to frequency. ANSI has used the value of 0.4 W/kg as the basis for its guide for safety.

At low frequencies the engineering calculations indicate a reduced coupling of energy into the human body. This has led ANSI and others to consider other criteria for the limitation of exposure at frequencies below 1 MHz. At frequencies down to 3 KHz ANSI has used induced current as the fundamental quantity to limit exposure. The basis for the limitation is (1) the prevention of radiofrequency burns and (2) the reduction of the probability of reactions caused by induced currents that exceed the thresholds for perception in adult males.

Precise definitions of dose which accurately reflect the level of potential harm requires an understanding of mechanisms. This is not to imply that it is necessary to understand mechanism to make an empirical assessment of a hazard or to set a protection standard. However, the process of defining dose and applying that definition to the establishment of a safety level requires assumptions that naturally exclude certain mechanisms. Generally this is a necessary simplification in the absence of a fundamental understanding of the interaction processes and the resulting detrimental (or beneficial biological response). For example, only a thermal mechanism can meet all of the criteria used in the ANSI 1982 standard even though no such assumption was made.

All of the historical approaches to defining dose discussed thus far, lead to simplified hazard assessment or dose response curve. An increase in the level of the dose or an increase in the dose rate results in an increase in the potential for causing a biological response and as the level is further increased the greater is the likelihood that such a response is harmful. Thus an increase in the time and/or intensity of exposure results in an increased risk. Furthermore, for the ANSI radiofrequency standard the dose level (SA or SAR) which may cause harm is independent of frequency.

Although we have no clear understanding of mechanisms of interaction with radiofrequency fields nor the mechanisms for any biological response, one can determine a list of important parameters to be considered for ELF exposure by considering some of the reported research dealing with the biological effect of ELF exposure. A few example ELF reports (see list of references) need to be considered. These include physical membrane responses (Marron et al., 1988), calcium dependent responses (Bawin & Adey, 1976; Blackman et al., 1980, 1985, 1988; Smith et al., 1978), effects on gene expression (Goodman et al., 1987, 1988, 1989; Philips & McChesney, 1989).

The motivation for undertaking some of these studies was not necessarily the

assessment of a potential hazard but rather an attempt to study electric field interactions with biological systems or to establish some basis for using electromagnetic fields for medical purposes. Thus these results do not necessarily indicate a hazard but describe a molecular or biological response that could possibly be used beneficially or detrimentally. Even though the data presented are, by and large, published, some of it is considered controversial and not replicated by independent investigators. However, these studies, at the very least, indicate the parameters (dose or affecting dose) that must be taken into consideration when conducting or analyzing research or epidemiological studies. The parameters to be considered are:

1. Frequency
2. Modulation
3. Intensity (minimum & maximum)
 - A) Magnetic Field
 - B) Electric Field
 - C) Electric and Magnetic Field
5. Duration
6. Intermittency or Distribution of Repeated Exposure Over Time

The ionizing radiation approach was intended to obtain means of predicting deleterious effects following a given exposure with specified characteristics. In the ELF bioeffects area several changes in cell function were demonstrated. However, in most instances, for example, calcium efflux effects, the biological significance, beneficial (?), deleterious (?) is not established. Not enough is known about interaction mechanisms to identify a single physical quantity, or a set of such quantities, which could be used to predict effects. Litovitz et al. (1989) has recently suggested that exposure to ELF fields affects the rate constant of certain biochemical processes (such as RNA transcription). This hypothesis provides a plausible explanation for the occurrence of time and amplitude windows, and removes the mystery of this complex response. Further exploration and experimentation of this type may provide a productive model useful for exposure assessment.

There are, however, two classes of phenomena induced by ELF exposure to which the ionizing radiation approach seems to be useful. One class are the responses of excitable cell membranes, resulting in an altered nerve or muscle cell function. It was established (WHO EHC 69, Magnetic Fields) that exposures that result in current densities below 10 mA/m² do not induce biological effects or result only in minor ones. At induced current densities between 10 and 100 mA/m² well established effects without a health significance (e.g., magnetophosphenes) occur and facilitation of healing of bone non-unions was reported. In the range between 100 and 1000 mA/m² the excitability of the nervous system changes, with possible deleterious consequences. At induced current densities above 1 A/m² extrasystoles, ventricular fibrillation, respiratory tetanus, and death may occur. These phenomena occur once a certain intensity (induced current density) threshold is reached; the duration of exposure does not play a significant role. Thus these phenomena are dose-rate dependent, and as all dose-rate dependent effects demonstrate a well defined threshold. Such data can (and are) used to predict untoward health consequences, and may serve as a basis for exposure limitation.

The second class of phenomena reminiscent of ionizing radiation effects consists of the suspected link between ELF exposure and cancer or congenital malformations. Ionizing radiation acts as an initiator and promoter of cancer. Thus the end result is the outcome of an interplay of statistical processes: DNA damage and repair, promotion, progression, and interactions between the genetic susceptibility and diverse environmental insults. The final outcome can be quantified in terms of probabilities of incidence of cancer or congenital malformations in exposed populations over a certain period of time. The dosimetric quantity used for predictive purpose is the mean exposure level (dose rate) per year (or the integral dose over one year) per man or rather per one individual in the exposed population. These stochastic effects have no demonstrable threshold, and the

prediction is meaningful only if the time factor is taken into account. All of the available epidemiological studies on cancer incidence in electromagnetic field exposed populations attempt to relate the incidence to exposure levels (dose rates). Because of the nature of the effects studied (i.e., their dependence on an interplay of statistical processes), which require a given length of time to become manifest, this approach cannot serve to prove or to disprove the association between exposure and cancer or congenital malformations. This applies irrespective of the mechanisms involved. One of the important points of my talk is the statement that further epidemiological or experimental studies on the association between ELF exposure and cancer or congenital malformations will not yield any useful results as long the duration of exposure is not taken into account.

For ELF exposure in general, however, the traditional concept of dose and dose rate is not applicable. The possible occurrence of time and intensity windows greatly complicate the problem. Furthermore, time must be considered on several different scales. Continuous exposures may be less effective in creating a response than intermittent exposures. The period of intermittence may be on the order of hours, seconds, or fractions of a second. The experimental results suggest that the exposure field may act more as an on/off switch rather than providing energy to cause an effect. This switch will not respond or over responds to a "large" field. The theoretical model proposed by Litovitz et al suggest effects on rate constants. In this case one can consider the field to be promoting or inhibiting a normal process. The amount of energy supplied by the field may be minute or immeasurable. The presence of the field is the important entity and not necessarily the total energy extracted from the field (dose).

REFERENCES

- American National Standards Institute, ANSI C95.1-1982: American National Standard safety levels with respect to human exposure to radiofrequency electromagnetic fields, 300 kHz to 100 GHz. IEEE. New York, N.Y.
- Bawin SM, Adey WR (1976): Sensitivity of calcium binding in cerebral tissue to weak environmental electric fields oscillating at low frequency. Proc Natl Acad Sci USA 73:1999-2003.
- Blackman CF, Benane SG, Joins WT, Hollis MA, and House DE (1980): Calcium-ion efflux from brain tissue: Power density versus internal field-intensity dependencies at 50-MHz RF radiation. Bioelectromagnetics 1:277-283
- Blackman CF, Benane SG, House DE, and Joins WT (1985): Effects of ELF (1-120 Hz) and modulated (50 Hz) RF fields on the efflux of calcium ions from brain tissue in vitro. Bioelectromagnetics 6:1-11
- Blackman CF, Benane SG, Rabinowitz JR, House DE, and Joins WT (1985): A role for the magnetic field in the radiation-induced efflux of calcium ions from brain tissue in vitro. Bioelectromagnetics 6:327-337
- Blackman CF, Benane SG, Elliott, DJ, House DE, and Pollock MM (1988): Influence of calcium on the efflux of calcium ions from brain tissue in vitro. Bioelectromagnetics 9:215-227
- Cooper MS, Miller JP, and Fraser SE (1989): Electrophoretic repatterning of charged cytoplasmic molecules within tissues coupled by gap junctions by externally applied electric fields. Developmental Biology 132:179-188
- Environmental Health Criteria No.16 (1987): Magnetic fields. World Health Organization, Geneva

Goodman R, Abbot J, and Henderson AS (1987): Transcriptional Patterns in the X chromosome of *Sciara coprophila* following exposure to magnetic fields. Bioelectromagnetics 8:1-7

Goodman R, and Henderson AS (1988): Exposure of salivary gland cells to low frequency electromagnetic fields alters polypeptide synthesis. Proc Natl Acad Sci 85:3928-3932

Goodman R, Wei L-X, Xu J-C, and Henderson A (1989): Relationship of transcript quantity to signal amplitudes. In Proceeding of the Bioelectromagnetics Society, 11th Annual Meeting, 18-22 June, Tucson, AZ, p. 29-30

International Commission on Radiation Units and Measurements (ICRU) 1979: Quantitative concepts and dosimetry in radiobiology. ICRU report 30, Washington, D.C.

Litovitz TA, Montrose CJ, Goodman R, Elson EC (1989): A mathematical model for the effect on cellular transcription of low frequency electromagnetic radiation. In Proceedings of the Bioelectromagnetics Society, 11th Annual Meeting, 18-22 June, Tucson, AZ, p. 49-50

Marron MT, Goodman EM, Sharp PT, Greenebaum B (1988): Low frequency electric and magnetic fields have different effects on the cell surface. FEBS Ltrs 230:13-16

National Council on Radiation Protection and Measurements (NCRP) 1981: Properties quantities and units, biophysical interactions and measurements. NCRP report 67, Washington, D.C.

Phillips JL, and McChesney L (1989): Effect of pulsed magnetic field exposure on cellular levels of RAS P21 in a human leukemia cell line. In Proceedings of the Bioelectromagnetics Society, 11th Annual Meeting, 18-22 June, Tucson, AZ, p. 9

Smith SD, McLeod BR, Liboff AR, and Cooksey K (1987): Calcium cyclotron resonance and diatom mobility. Bioelectromagnetics 8:215-227

Thomas JR, Schrot J, and Liboff AR (1986): Low intensity magnetic fields alter operant behavior in rats. Bioelectromagnetics 7:349-357

RELATIVE POTENCY CAN BE USED TO EVALUATE ELF HEALTH RISKS

Clay E. Easterly
Oak Ridge National Laboratory
Oak Ridge, Tennessee

Operated by Martin Marietta Energy Systems, Inc.
under contract DE-AC05-84OR21400 with the U.S. Department of Energy.

INTRODUCTION

Approximately one decade ago, we at ORNL began to be concerned about the possible adverse biological effects that could accompany exposure to non-ionizing electromagnetic fields. Initially, interest was centered around the quasi static magnetic fields of fusion reactors (Easterly 1982); soon, questions about power frequency fields were seen to be qualitatively similar (Easterly 1981). Laboratory studies revealed an alarming array of results, often conflicting and without consistency. Our particular interest at ORNL was in how to view the body of available data in a manner that allowed a meaningful interpretation of the hazard, or perhaps even health risk, from exposure to the nearly ubiquitous fields.

A review of the available data showed that reproducible biological effects were difficult to identify. Neither in vivo or in vitro experiments were able to provide a consistent level of effect. Consequently, the challenge to evaluate the level of anticipated hazard, or risk, was clear. There was neither a mechanistic understanding, nor a consistent phenomenological outcome at our disposal. Normally, in performing a risk evaluation, one or the other of these two types of information is imperative. Lacking either of these, we began to urge the Department of Energy to initiate or stimulate epidemiologic research. Eventually, other agencies and organizations saw this same need and have begun this work.

Notwithstanding the present and planned epidemiological research, there is a vast and varied body of research using power frequency (or some surrogate) electromagnetic fields. Can one get a net around the information in order to provide more than a qualitative perspective? This question, which is still before us, was present in the early 1980s. At that time, we proposed two methods to better address the question. I contend now, as I did several years ago, that these methods are improvements to existing techniques; progress has been made to support these assertions. The two quite different approaches to which I refer are embodied in the terms **meta-analysis** and **relative potency**. The first of these will be discussed later in these proceedings by Dr. Kay Kimball, who worked with us in developing our ideas on meta-analysis. The second, relative potency, will be the focus of the balance of my remarks.

RELATIVE POTENCY BACKGROUND

Relative potency is not new in any sense of the word. It has been around for millennia. Relative comparisons are ubiquitous in science and engineering. The roots of the present application are in pharmacology wherein drugs are compared in terms of relative doses required to produce the same therapeutic action. For example, if it takes two aspirin tablets to cure a headache and one tablet of a new drug, the relative potency of the new drug would be 2. That is, the new drug is twice as potent towards relieving headaches as the reference headache medicine, aspirin. The relative potency framework is used to compare doses required to produce equal levels of effects. These relative comparisons in pharmacological uses are performed without making assumptions about biological mechanisms. The approach being pursued at ORNL is also based on observed

biological response rather than hypothesized mechanism of action. The scale is constructed by expressing results as a function of an arbitrarily selected standard, usually a well studied toxic agent. A distinguishing feature is that it is "data intensive and model sparse," as such it incorporates a broad spectrum of biological data. Thus, rather than focusing on identification of the single "best" animal experiment, as is done in "quantitative risk assessment" (EPA 1989), an attempt is made to utilize all available sources of data.

We introduced such an approach to our sponsor several years ago for application to the question of ELF electromagnetic fields. The concept received a lengthy review from a variety of recognized experts in the ELF research arena. Skepticism was voiced about the utility of such an approach that does not incorporate an understanding of mechanisms. None of the reviewers seemed to remember that the entire body of radiation protection is based on relative potencies of the various types of ionizing radiations derived from exposure-effect data rather than mechanistic models. Given the response by the review panel, our next move was to demonstrate that a relative potency framework, as we envisioned it, could operate successfully without a detailed knowledge of mechanisms.

The framework used has undergone continuous development in the Health Effects and Epidemiology Group at the Oak Ridge National Laboratory for approximately a decade (Griffin et al. 1979; Walsh et al. 1982; Dudney et al. 1983; Jones et al. 1985, 1988; Glass, 1987). The present approach is intended to provide an alternative to the current standard methods for estimating carcinogenic potency (see, for example, Anderson 1983; EPA 1989); it is, hopefully, a more scientifically defensible framework for evaluating potencies of a wide range of insults. The greater potential for defensibility of the ORNL approach derives from different philosophic bases relative to data selection and low dose extrapolation. The results of this work will demonstrate that this ranking can take place without explicit knowledge of mechanism of action; therefore, the use of relative potency in assessing power frequency health risks is not bound by our lack of mechanistic knowledge.

CONCEPTUAL PERSPECTIVE

The thought process of how to address the issue of mechanisms is as follows. Chemical carcinogens, those chemicals recognized by the International Agency for Research on Cancer (IARC) as human or potential human carcinogens, represent a wide array of mechanisms of action. That is, the means by which these chemicals produce cancer in animal systems may be quite different from one chemical to another. If, by some technique, these carcinogens could be ranked relative to a reference compound of one mechanism of action, that ranking could be compared with a ranking on the basis of another reference compound that had a different mechanism of action. This process could be carried out for reference compounds encompassing the known classes of mechanism of action. If the rankings are dependent on mechanism, then the potency rankings based upon the different reference compounds would be different. If the rankings are independent of mechanism of action, then the technique could serve as a valuable model for future studies. As will be described later, the results of the research effort demonstrated that, in fact, the approach developed at ORNL, using a wide spectrum of long-term whole animal data is useful in providing a ranking of carcinogenic potency independent of mechanism of action for the reference compound (Glass 1987).

The most appropriate next study would be one that allowed for the greatest diversity in biological responses to demonstrate the application of the relative potency approach to a diverse set of bioeffects endpoints such as are present in the ELF bioeffects database. The study undertaken was to determine if potency ranking as predicted with long-term whole animals could be replicated using short-term tests. The results of that study (Glass 1987) demonstrate that a potency ranking on the basis of short-term tests provides essentially the same ranking of the carcinogens as the long-term whole animal studies (Glass 1987).

APPLICATION OF CONCEPT TO DATA

The relative potency framework as utilized at ORNL involves a comparison of dose levels of different agents (chemicals) to give a common outcome in a common experimental protocol. Thus, if a variety of experimental protocols are available in the literature for a given test chemical and a reference chemical, a range of these ratios is obtained. These ratios represent a distribution in sensitivity of the various test systems, which themselves represent a variety of species, strains, methods of dosing, time course of dosing, etc. Because nearly all of the biological test systems are "tuned" to be sensitive to some insult, they should not be thought of as individually representing humans as a whole organism. Rather, they can be thought of as collectively representing the distribution of responses by living organisms.

In a general sense, lacking evidence to the contrary, no single test system should be thought of as contributing more or less information than any other; they each represent a response of a specific organism to a specific exposure (Glass 1987). Therefore, all data are considered as providing equivalently useful information. Not having information upon which to determine the underlying distribution of responses, the summary measure determined for use in this work is the median (Jones et al. 1985). When determining the relative potency of a test chemical with respect to that of a well studied reference chemical for a variety of experimental conditions, a range of ratios is obtained on the basis of comparing doses required to produce the same endpoint in the same experimental conditions. The relative potency is then the median of the range of ratios.

For evaluation of these concepts, 52 carcinogens were selected, mostly, from the list of carcinogens published by the IARC. Long-term whole animal data on tumorigenicity and short-term data were obtained from the Registry of Toxic Effects of Chemical Substances (RTECS). To avoid any potential bias, 13 chemicals from among the 52 were chosen to serve as reference compounds. These reference compounds were chosen to represent different mechanisms of action.

In determining whether or not mechanism of action was important for ranking the long-term whole animal tumorigenic studies, 13 rankings of the 52 chemicals were performed. These 13 lists were compared with one another, two-by-two, for consistency using the Spearman Rank Correlation test. Nearly all of the 78 pairs that were ranked had Spearman Rho values of 0.8 or greater and p-values were typically less than 0.0001 (Glass 1987). These results provide a high degree of confidence that the relative potency framework, as used, can be expected to reliably rank carcinogenic potency regardless of the mechanism of action of the agents being ranked.

Short-term tests that involve some type of mutation are often thought to be indicative of carcinogenic potential. However, the tests are most often performed on cellular systems that do not have the capability to respond like whole animals. Furthermore, in order that they can be sensitive to a specific type of insult, they are almost always mutated systems to begin with. These tests involve bacteria, yeast, mammalian cells and sometimes, host animals. It would be stretching the imagination to suggest that any one of these test systems was predictive of or, mechanistically reflected the response of any other test system, let alone the wild type whole animal system represented by the human. However, when relative potencies using data from short-term tests are evaluated for a test chemical with respect to a well studied reference chemical, a distribution of ratios is found similar to that for the whole animal tumorigenicity data. Summary relative potencies are again described by the median value. Using short-term data, the 13 reference chemicals were used to rank the 52 carcinogens, just as before. As might be expected, the range of responses (ratios) within the short-term assays was greater than within the whole-animal assays and therefore the median was less stable. As a consequence, when the 13 rankings of the short-term tests were compared with the 13 rankings of the tumorigenic tests using the Spearman Rank Correlation test, the

correlations were lower than before, typically 0.4, and the p-values higher, typically 0.01 to 0.05 (Glass 1987). Considering the fact that the individual ratios for a given test and reference chemical often vary by factors of 1,000,000 or more, the consistency of Rho (which can vary from -1 to +1) that generally varied between 0.35 and 0.6 is remarkable. Only three p-values were greater than 0.05; they were 0.16, 0.08, and 0.06. Again, given the consistency of results and the summary statistic (tested against a two-tailed distribution), the strength of the correlation is remarkable.

PERSPECTIVE ON APPLICATION TO ELF ELECTROMAGNETIC FIELDS

On the basis of the similar ranking of the carcinogenic potency of the chemicals between whole animal data and the widely varying short-term data, a cautious step can be made toward analyzing effects derived by ELF electromagnetic fields, and for a variety of different biological endpoints. The variation in the biological mechanisms between the different chemicals and the long-term and short-term tests should encompass the differences in biological action produced by ELF fields and potential chemical reference agents. This being the case, it is believed that the relative potency framework can be cautiously applied to the ELF data base to help in a variety of ways. It can be of assistance in determining the (carcinogenic) potency of ELF fields relative to other agents. Such an evaluation is valuable in designing epidemiologic studies. It is also believed that application of this work can be of assistance to the laboratory research effort by providing a method for a collective interpretation of a variety of diverse experimental procedures.

REFERENCES

- Anderson, E.L. (1983). Quantitative approaches in use to assess cancer risk. Risk Anal. 3(4):277-295.
- Dudney, C.S., Walsh, P.J., Jones, T.D., Calle, E.E., and Griffin, G.D. (1983). On the use of relative toxicity for risk estimation. pp 499- 513 in Short-term Bioassays in the Analysis of Complex Environmental Mixtures III. M.D. Waters, S.S. Sandhu, J.Lewtas, L. Claxton and S.N. Nesnow, eds., Plenum Press, New York.
- Easterly, C.E. (1982). Cardiovascular Risk from Exposure to Static Magnetic Fields. Am. Ind. Hygiene Assn. 43:533-539.
- Easterly, C.E. (1981). Cancer Link to Magnetic Field Exposure: A Hypothesis. Am. J. Epidemiol. 114(2):169-174.
- Glass, L.R. (1987). Use of short-term test systems for the prediction of the hazard represented by potential chemical carcinogens. Ph.D. Dissertation, the University of Texas Health Science Center at Houston, School of Public Health, Houston, Texas.
- Griffin, G.D., Jones, T.D., and Walsh, P.J. (1979). Chemical cytotoxicity: pp 723-732 in Polynuclear Aromatic Hydrocarbons, P.W. Jones and P. Leber, eds., Ann Arbor Science Publishers, Inc., Ann Arbor, Mich.
- Jones, T.D., Walsh, P.J., and Zeighami, E.A. (1985). Permissible concentrations of chemicals in air and water derived from RTECS entries: a "RASH" chemical scoring system. Tox. Ind. Health 1(4):213-234.
- Jones, T.D., Walsh, P.J., Watson, A.P., Owen, B.A., Barnthouse, L.W., and Sanders, D.A. (1988). Chemical scoring by a Rapid Screening of Hazard (RASH) method. Risk Anal. 8(1):99-118.

Walsh, P.J., Jones, T.D., Griffin, G.D., Dudney, C.S., Calle, E.E., and Easterly, C.E. (1982). Risk assessment approaches: General definitions, limitations, and research needs. J. Environ. Sci. Health A17:541-552.

EPA (U.S. Environmental Protection Agency) (1989). Workshop Report on EPA Guidelines for Carcinogenic Risk Assessment, EPA/625/3-89-015, March 1989, Washington D.C. 20460.

DISCUSSION Special Problems I

Peter G. Groer
Oak Ridge Associated Universities
Oak Ridge, Tennessee

The first paper in this session, "What Is Dose?" by Mays Swicord discusses the question of "dose" and dose assessment for ELF fields. He draws comparisons between ionizing and non-ionizing radiations and makes the important point that choice of the appropriate dose measure presupposes an understanding of the mechanism that produces an harmful effect in an exposed individual. Despite the fact that ionizing photons (x-rays and gamma rays) and ELF fields are both electromagnetic radiation, it is clear that for energetic reasons, ELF fields interact differently with living matter. They cannot break molecular bonds and they interact instead by, for instance, inducing electric currents or by acting directly on moving charged particles in tissues. Another important difference between ionizing and non-ionizing radiations besides the lower energy is the simultaneous presence of the electric and the magnetic field vectors that interact individually in different fashions with exposed tissue. ELF fields interfere through a great number of different mechanisms with cells, the cellular surface, biological macromolecules, and electrolytes, thus affecting normal biological processes. This influence on ongoing biological processes is the reason for the macro- and microdosimetric complexity of ELF fields and also of electromagnetic radiations of higher frequency. It seems, therefore, extremely unlikely that a single "dose" measure could be found that would suffice as a quantitative description of all possible ELF-interferences with normal biological processes. In the future, painstaking modeling of biological processes with different dose measures seems to be necessary to unravel the induction of carcinogenic or of other deleterious health effects by exposure to ELF fields. This paper makes this clear directly or by implication.

The second paper entitled "Relative Potency Can Be Used to Evaluate ELF Health Risks," by Clay Easterly discussed possible application of a relative potency approach to evaluate the risks from exposure to ELF fields. The author described the historical development of the underlying ideas at Oak Ridge National Laboratory and thought that a relative potency framework, which was applied to chemical carcinogens, could also be "cautiously applied to the ELF data base."

MULTIPLE COMPARISONS IN STATISTICAL ANALYSIS

Abdelmonem A. Afifi
UCLA School of Public Health
Los Angeles, California

Introduction.

In most statistical applications, the investigator typically makes several inferences with each subject to a certain probability of error. There is also a joint probability of error associated with the total set of inferences made. The situation can be formulated in general terms as follows. Let $S_1, S_2, \dots, S_q$ be q statements (inferences) to be made by the investigator. For each S_i we attach a probability α_i of being in error, i.e., $\alpha_i = P(S_i \text{ is incorrect})$ or $1 - \alpha_i = P(S_i \text{ is correct})$. The joint probability of error is α , i.e., $\alpha = P(\text{some } S_i \text{ are incorrect})$ or $1 - \alpha = P(S_1, \dots, S_q \text{ are all correct})$.

The problem of simultaneous inference (or multiple comparisons) is how to select the best procedure for making the statements in such a way that $\alpha_1, \dots, \alpha_q$ and α are specified quantities. For example, S_i may be a confidence interval for the i th parameter, $i = 1, \dots, q$. The problem is how to construct individual confidence intervals with confidence levels $(1 - \alpha_i)$, so that the joint confidence level is $(1 - \alpha)$. For another example, let H_i , $i = 1, \dots, q$ be individual null hypotheses and $H = \cap_{i=1}^q H_i$ be the overall null hypothesis to be tested. The question is how to test the individual H_i , with significance levels α_i , so that the joint significance level associated with H is a specified quantity α .

This paper is not intended to present a comprehensive treatment of the subject. For such a treatment, the reader is referred to the books by Miller (1981) and Hochberg and Tamhane (1987), and any of several texts on experimental design, e.g., Fleiss (1987). The remainder of this paper discusses in some detail two statistical methods relevant to electromagnetic field (EMF) research, namely, linear regression and logistic regression. Three other references are mentioned as examples of related recent results. The paper concludes with general recommendations regarding how and when to perform multiple comparisons. Although numerous multiple comparisons methods have been derived in given situations, this paper will concentrate on only two of them in order to emphasize the basic concepts.

Linear Regression.

Simple linear regression is a convenient framework for understanding some fundamental ideas. Each observation y_i can be written as

$$y_i = \beta_1 + \beta_2 x_i + e_i, \quad i = 1, \dots, n,$$

where the e_i are assumed to be independent $N(0, \sigma^2)$ error terms. The least squares estimates $\hat{\beta}_1$, $\hat{\beta}_2$, their estimated standard errors, $s.e.(\hat{\beta}_i)$, $i = 1, 2$, and the unbiased estimate s^2 of σ^2 , can be easily obtained. The usual individual confidence intervals (CI's) for β_i can be written as

$$\beta_i \in \hat{\beta}_i \pm s.e.(\hat{\beta}_i) t_{\frac{\alpha}{2}; n-2}, \quad i = 1, 2,$$

using standard notation. The problem is that the joint confidence level for β_1 and β_2 is $< 1 - \alpha$. Schöffé (1953) developed a general multiple comparisons method giving the

following CI's

$$\beta_i \in \hat{\beta}_i \pm s.e.(\hat{\beta}_i)[2F_{\alpha;2,n-2}]^{1/2}, \quad i = 1, 2.$$

where F_α is the upper α quantile of the F distribution. The theory developed by Scheffé guarantees that the joint confidence level is $1 - \alpha$. Furthermore, the method can be used to construct CI's for any linear combination of β_1 and β_2 , while guaranteeing a joint confidence level $1 - \alpha$ for this infinite number of CI's. We are also assured that the overall null hypothesis

$$H : \beta_1 = 0 \text{ and } \beta_2 = 0$$

is rejected if and only if at least one such CI excludes zero. This impressive generality, however, is achieved at the cost of providing CI's that are wider than the individual CI's.

An alternative method is based on the Bonferroni inequality. This inequality states that

$$\alpha \leq \sum_{i=1}^q \alpha_i \quad \text{or} \quad 1 - \alpha \geq 1 - \sum_{i=1}^q \alpha_i.$$

The Bonferroni CI's

$$\beta_i \in \hat{\beta}_i \pm s.e.(\hat{\beta}_i)t_{\frac{\alpha}{2};n-2}, \quad i = 1, 2,$$

each with confidence level $(1 - \alpha/2)$, thus guarantee that the joint confidence level for β_1 and β_2 is at least $1 - \alpha$. They are also shorter than the Scheffé CI's for β_1 and β_2 . However, they assume that these two CI's are the only statements to be made.

The Scheffé method can also be used to construct a confidence band for the whole regression line. The least squares estimate of $\mu_x = \beta_1 + \beta_2 x$ is $\hat{\mu}_x = \hat{\beta}_1 + \hat{\beta}_2 x$ and its estimated s.e. is $s.e.(\hat{\mu}_x) = s \left[\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right]^{1/2}$. The joint confidence level for the CI's $\mu_x \in \hat{\mu}_x \pm s.e.(\hat{\mu}_x)[F_{\alpha;2,n-2}]^{1/2}$ for all x is $1 - \alpha$. Similarly, the predicted value y for a

given x is $\hat{y}_x = \hat{\beta}_1 + \hat{\beta}_2 x$, and its estimated s.e. is $s.e.(\hat{y}_x) = s[1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum (x_i - \bar{x})^2}]^{1/2}$. The predictive intervals (PI's) $y_x \in \hat{y}_x \pm s.e.(\hat{y}_x)[2F_{\alpha;2,n-2}]^{1/2}$ have a joint confidence level $1 - \alpha$ for all x . If the investigator is interested in only q specific values of x , then Bonferroni CI's and PI's can be computed by substituting $t_{\frac{\alpha}{2q};n-2}$ for $[2F_{\alpha;2,n-2}]^{1/2}$ in the above formulas. We note that if q is large, then the Scheffé intervals are shorter than Bonferroni's, and *vice versa* (see Miller, 1981).

Both methods are applicable to the general linear model. For example, in the multiple linear regression model

$$y = \beta_1 + \beta_2 x_1 + \dots + \beta_p x_{p-1} + e = \beta'X + e$$

we can construct simultaneous CI's for $\mu_x = \beta'X$. Denote by $\hat{\mu}_x$ the least squares estimate of μ_x and by $s.e.(\hat{\mu}_x)$ its estimated standard error. The two methods then give the following CI's:

Scheffé CI (for all values of X):

$$\mu_x \in \hat{\mu}_x \pm s.e.(\hat{\mu}_x)[pF_{\alpha;p,n-p}]^{1/2},$$

Bonferroni CI (for q specified values of X):

$$\mu_x \in \hat{\mu}_x \pm s.e.(\hat{\mu}_x)t_{\frac{\alpha}{2q};n-p}.$$

Logistic Regression.

The method of logistic regression has gained great popularity in epidemiological studies, partly because of its validity in case-control studies. Suppose that E is an event of interest (e.g., the occurrence of a given disease in a given period of time) and X is a vector

of $\mathbf{P}$ covariates (confounders, effect modifiers, etc.) The model equation can be written as:

$$P(E | \mathbf{X}) = \frac{1}{1 + e^{-\beta' \mathbf{X}}}.$$

Many packaged programs can be used to compute the maximum likelihood estimates $\hat{\beta}_i$ of β_i and their estimated asymptotic standard errors $s.e.(\hat{\beta}_i)$, $i = 1, \dots, p$. The theory of Scheffé can be used to produce approximate CI's with a joint asymptotic confidence level $1 - \alpha$, namely,

$$\beta_i \in \hat{\beta}_i \pm s.e.(\hat{\beta}_i)[\chi_{\alpha;p}^2]^{1/2}, \quad i = 1, \dots, p,$$

and Bonferroni CI's with a minimum asymptotic confidence level of $1 - \alpha$ are:

$$\beta_i \in \hat{\beta}_i \pm s.e.(\hat{\beta}_i)Z_{\frac{\alpha}{2p}}, \quad i = 1, \dots, p,$$

where Z_u and χ_u^2 are the upper u quantiles of the standard normal and χ^2 distributions, respectively.

If the sample available is from a cohort study, then we can construct CI's for the probability $P(E | \mathbf{X})$ based on a given vector $\mathbf{X}$. Let Σ be the covariance matrix of the estimated vector $\hat{\beta}$. Some packages, e.g. GLIM, produce an asymptotic estimate $\hat{\Sigma}$ of Σ . Letting $\hat{\beta}' \mathbf{X}$ be the estimate of $\beta' \mathbf{X}$, and $\sigma_x^2 = \text{Var}(\hat{\beta}' \mathbf{X}) = \mathbf{X}' \Sigma \mathbf{X}$ then $\hat{\sigma}_x^2 =$ estimate of $\sigma_x^2 = \mathbf{X}' \hat{\Sigma} \mathbf{X}$.

With this information, we can compute simultaneous CI's for $\beta' \mathbf{X}$ as follows.

Scheffé-type CI (for all $\mathbf{X}$):

$$(L_1, L_2) = \hat{\beta}' \mathbf{X} \pm \hat{\sigma}_x [\chi_{\alpha;p}^2]^{1/2},$$

Bonferroni-type (for q preselected $\mathbf{X}$'s):

$$(L_1, L_2) = \hat{\beta}' \mathbf{X} \pm \hat{\sigma}_x Z_{\frac{\alpha}{2q}}.$$

For either method, we can construct the CI's for $P(E | X)$ as:

$$\text{lower confidence limit} = \frac{1}{1 + e^{-L_2}}, \text{ upper confidence limit} = \frac{1}{1 + e^{-L_1}}.$$

If the sample is obtained from a case-control study, then the estimate of $P(E | X)$ is not consistent. Bock and Afifi (1986) discuss various methods for adjusting the estimate to produce consistency. The adjusted estimate can then be used to compute CI's.

Other Recent Developments.

The area of multiple comparisons continues to attract many researchers. We selected three papers to present as examples of recent developments in statistical areas related to research in EMF effects. Afifi et al (1986) compared various methods for computing CI's for survival probabilities based on censored data. They recommend using a Bonferroni-type method for most of the situations they examined. Dixon and Divine (1987) considered the Cox proportional hazards model, in which the coefficients for indicator variables represent relative risks. They presented a method which sometimes results in shorter CI's than Bonferroni's. Fairweather (1987) examined studies in which cases are compared to several control groups in order to increase the validity of the results. Requiring significance versus each control group can be too stringent. He, therefore, derived a more powerful maximum range test that compares cases with all control groups simultaneously.

General Recommendations.

The investigator looking for a simple general method of multiple comparisons would do well by selecting the Bonferroni procedure. It is applicable in a very wide variety of situations. For confidence intervals, one would construct the usual one-at-a-time CI's with a single modification: *divide α by q* , the number of CI's computed. For hypothesis testing,

one would compute the usual one-at-a-time p -values and then: *multiply each p -value by q* , the number of tests performed. In any given situation, the investigator is of course advised to use a sharper method if such exists. A sharp method is one that gives shorter CI's or more powerful tests. We note, in particular, that Fuchs and Sampson (1987) describe methods to sharpen the Bonferroni approximation in the general linear model.

It is not as easy, however, to make a general recommendation regarding the questions: when should multiple comparisons be performed, or what should the number of statements q be? In one extreme, we can insist that all statistical inferences be considered simultaneously. In this scenario we may never have a significant result, since q is virtually infinite. In the other extreme, we can examine each statistical statement in isolation of all others. This can also be inappropriate in certain situations. For example, in the one-way analysis of variance we should adopt a simultaneous inference attitude. If the overall hypothesis of no effect is accepted, we should quit. If the hypothesis is rejected, we could, and perhaps should, proceed with an appropriate multiple-comparisons method to examine a set of potentially responsible contrasts.

A principle is needed to guide us in the vast area lying between the above-mentioned extremes. Miller (1981) presents such a guide for experiments from which conclusions will be drawn and reported:

The natural family [of statements] for the author in the majority of instances is the individual experiment of a single researcher.

More generally, statistics can be defined as the art and science of making decisions in the face of uncertainty. In the final analysis, it is the weight of the evidence that will carry

the day. Moreover, in the EMF area, final decisions will probably be made by government regulators and the courts of law. The statistical analysis should, therefore, be aimed at making the most convincing case possible, given the available evidence.

References.

- Affi, A.A., Elashoff, R.M. and Lee, J.K. (1986). "Simultaneous nonparametric confidence intervals for survival probabilities from censored data." *Stat. Med.*, 5:653-62.
- Bock, J. and Affi, A.A. (1988). "Estimation of probabilities using the logistic model in retrospective studies." *Computers and Biomedical Research*, 21:449-470.
- Dixon, D.O. and Divine, G.W. (1987). "Multiple comparisons for relative risk regression: Extension of the K ratio method," *Stat. Med.*, 6:91-7.
- Fairweather, W.R. (1987). "Comparing proportions exposed in case control studies using several control groups," *Am. J. Epidemiol.*, 126:170-8.
- Fleiss, J.L. (1986). The design and analysis of clinical experiments. Wiley, Interscience, New York.
- Fuchs, C. and Sampson, A.R. (1987). "Simultaneous confidence intervals for the general linear model," *Biometrics*, 43: 457-69.
- Hochberg, Y. and Tamhane, A.C. (1987). Multiple comparison procedures. Wiley, New York.
- Miller, R.G. (1981). Simultaneous statistical inference, 2nd ed. Springer-Verlag, New York.
- Scheffé, H. (1953). "A method for judging all contrasts in the analysis of variance." *Biometrika*, 40:87-104.

INVESTIGATING CELLULAR RESPONSES TO EMF: CAN STATISTICS HELP?

Anthony D. Thrall
Electric Power Research Institute
Palo Alto, California

The public's general concern that electric or magnetic fields (EMF) may be harmful has motivated a wide variety of scientific studies. One of the more exploratory lines of scientific inquiry is the investigation of cellular responses to EMF. If such responses are substantiated, they may help to refine hypotheses about how EMF could induce or promote cancer.

Several scientists are investigating cellular responses to EMF. I will confine my remarks, however, to the research of Drs. Reba Goodman, Martin Blank, and their colleagues for it is this work that is partially supported by Electric Power Research Institute (EPRI), and with which I am therefore becoming most familiar. I must also confine my remarks to prospects for statistical support of this research, but I am confident that many of these prospects will soon be realized.

The questions I would like to discuss are whether and how statisticians can help in this investigation, or more generally, in such exploratory scientific investigations. Leaders of the statistical profession have urged the rest of us to become "whole" (Box, 1979; and Tukey, 1979), that is, for each of us to be capable of providing data analysis, probabilistic modeling, or statistical inference, as needed. Which if any of these activities can assist the investigation of cellular responses to EMF?

STATISTICS IN SCIENTIFIC RESEARCH

In his opening remarks, Dr. Ginevan distinguished two goals of the research being conducted on the health effects of EMF: (1) to increase scientific understanding; and (2) to inform public policy. My discussion concerns the former goal, i.e., whether and how statistics can assist the scientific investigation of rather broad hypotheses. I would like to begin the discussion by presenting a view of the role of statistics in scientific research.

It seems useful, first of all, to distinguish three phases or modes of scientific inquiry: (1) conjecture, (2) investigation, and (3) criticism. Conjecture consists of generating various hypotheses about the phenomena under consideration. For example, Ross Adey and others have conjectured that cyclotron resonance is the mechanism of cellular responses to EMF at extremely low intensities. Investigation consists of an experiment or an observational study to test one or more hypotheses, examples of which I will mention in a moment. And by criticism I mean critical reflection on the degree to which scientific evidence corroborates a hypothesis or favors it over competing hypotheses.

As children we were taught that scientific inquiry starts with theory, proceeds to investigation, and ends with criticism (or reevaluation). I think most of us would agree, however, that scientific inquiry cycles through all three modes, and that any one of the three modes can lead to the other two. This is represented in Figure 1.

Within this cycle of inquiry are several tasks (also shown in Figure 1) in which statistics can offer help: (1) interpreting results, (2) formulating models, (3) designing investigations, and (4) analyzing data.

I think it is also worth mentioning that several motive forces are required to make the cycle go. The principal investigator's motivations include scientific curiosity and the desire to professional advancement. A funding organization such as the NSF may want to

advance scientific knowledge in an area that somehow seems important. Other funding organizations, such as the EPA or an industrial association, may be seeking the resolution of a pressing social issue. And concerns and aspirations of society, as expressed or influenced by the media and legislative representatives, also deserve recognition as a distinct but related form of motivation.

In particular, these different parties to the scientific inquiry have their distinct goals, which may be conjoined temporarily but will eventually part. I think these separate goals deserve both recognition and respect by those developing a strategy of investigation.

CELLULAR RESPONSES TO EMF

Early evidence that EMF might alter basic cellular functions came from the observation that certain weak, pulsing, electromagnetic fields (PEMFs) helped to heal bone fractures that were unresponsive to other treatments (see reference in Goodman, Bassett, and Henderson, 1983). Subsequent work by Goodman and her colleagues suggests that these and other wave-forms induce RNA transcription and alter the statistical distribution of cell proteins. Blank and Goodman (1989) hypothesize that various actions on the cell, including thermal stress and EMF, interrupt the cell's biosynthesis of proteins. Blank has developed an electrochemical model that accounts for the hypothesized interruption.

The evidence is now fairly well substantiated that weak PEMFs having special wave-forms do aid bone fusion. For example, Lee (1989) reports that among a total of 76 patients in a double-blind, multi-center study, the rate of successful lumbar fusion was roughly 90% within the treatment group, as compared to only 70% or so within the control group, a statistically significant difference.

What are the cellular mechanisms by which PEMFs aid bone fusion? Why are some wave-forms more beneficial than others? These are subjects of ongoing investigations. It appears that these signals effect cellular transcription and the relative distribution of polypeptides (Goodman and Henderson, 1989).

The questions concerning cancer are different. First, are observed cellular responses to EMF a general response produced by other stresses? And do these cellular responses suggest that EMF may be a carcinogen? The latter question is the one most relevant to the EMF controversy, and we have a long way to go before we will be able to answer it.

Currently, then, bone healing is a demonstrated effect of EMF, whereas cancer is a suspected effect. Present investigations of the action of EMF on the cell advance along both lines of inquiry, perhaps because we still know so little. But the day may come when these two lines of inquiry will part.

STATISTICIAN AS CO-CONSPIRATOR

One way in which a statistician can contribute to an investigation is to join the investigative team, to analyze data, formulate probability models, or help to design the investigation.

Data Analysis

Data analysis is perhaps best characterized as an attitude, (i.e., a disposition toward entertaining many conjectures simultaneously and looking for clues in departures of the data from conjectured models). Data analysis is also an expanding set of techniques to summarize the data and stimulate conjecture about the data.

Tukey (1986) gives a concise example of data analysis. He notes that counts of radioactive disintegrations in successive time intervals led to a Poisson process model of the disintegrations. The model implies that a certain transformation of the histogram of

counts should approximate a straight line. This is in fact the case, which thus supports the Poisson hypothesis. But this view of the data does not reveal how well the data conform to the hypothesis. To illustrate what he means by data analysis, Tukey provides an alternative view of the data, namely the aforementioned linearized histogram minus the best fitting line. This plot of residuals (transformed data minus fitted model) enables us to assess the agreement between data and model, both visually and statistically. Moreover the plot may provoke useful questions about the residuals, (e.g., the degree to which they are due to measurement error, sampling error, or an overly simplistic model).

Models of Random and Systematic Variation

An activity that may be undervalued, both by the providers and users of statistical services, is the formulation of probability models. These models should, at a minimum, identify sources of variation and the components that vary independently of the remaining components so that the magnitude of variance components can be estimated. More ambitious models, (i.e., linear or nonlinear regression models, additionally specify the pattern of variation).

Of course, these models should also be interpretable by the principal scientific investigator. Hopefully they will also be provocative, i.e., will provoke useful thought when compared to the data.

Planning

One of the more widely acknowledged forms of statistical support is the design of experiments. This proceeds as follows:

- formulate questions and assumptions;
- identify alternative experimental designs
- compare the cost and expected information from alternative designs
- choose a cost-effective, informative design

From the time of Fisher's pioneering work in designing agricultural experiments at Rothamsted, statisticians have proved their value in this manner.

But there is another, less technical way in which statisticians can contribute to the planning of long-term investigations. That is, by facilitating discussion of the broader questions of research strategy.

- What are we trying to achieve?
- What are some alternative strategies?
- What are the advantages & disadvantages of these alternatives?

STATISTICIAN AS SYMPATHETIC SKEPTIC

If statisticians are known to some investigators to be helpful during the design stage of an experiment, they are known to a much broader set of investigators to be adjudicators of the experimental evidence supporting or negating a model, once the experiment is over.

I think a more useful perspective of this function of the statistician is that of a (hopefully friendly) skeptic. Of course, the principal investigator (PI) is herself or himself also a skeptic, as are the PI's colleagues both within the investigative team and outside the investigative team, where their views may be expressed as anonymous referees of papers submitted for publication.

But there are some advantages to asking a statistician to serve as skeptic, particularly when a long-term working relationship has been established. First, the statistician is not a competitor. Second, the statistician is often sufficiently acquainted

with the work to make helpful comments, yet not so intimately involved that skepticism is

impossible. Third, because the statistician is typically not an expert in the substantive field of inquiry, she or he requires explanations that are understandable to the broader community of scientists, which fosters interdisciplinary communication.

HOW MUCH CAN STATISTICIANS HELP?

I think there are several ways in which statisticians can aid the investigation of cellular responses to EMF. First, statistical analysis of experimental research can help to confirm that EMF does affect the cell, and can thus help to move current analyses from a qualitative stage to a quantitative stage.

Second, statistics can help to design and analyze the exploration of difference in cellular responses to different stimuli, (i.e., heat or other stresses, or various EMF frequencies, amplitudes, and wave-forms). (The latter may be most usefully probed by examining the frequency spectrum, [i.e., harmonic content, of the various wave-forms]).

Finally, statisticians may be able to contribute to the overall strategy for this line of inquiry. In particular, we need to distinguish the types of studies that are narrowly focused on whether EMF is a carcinogen, from studies that may make other scientific contributions. That is, we would do well to design studies that will advance the scientific understanding of cellular responses to stimuli, even if a decade from now EMF has been set aside as an environmental hazard.

More generally, how much can statistics help in other exploratory studies? I think the principal investigators should be encouraged to use statistics to confirm and quantify phenomena, design experiments that yield as much information as possible within practical limitations, and explore alternative explanations of experimental results.

Of course, the answer to just how useful statisticians can be is in our hands, more than anyone else's. We need to join the computer revolution, in particular to free ourselves of the confines of convenient textbook statistical models. We statisticians also need to consult with one another more, to take advantage of specialized expertise. This can be facilitated by putting directories of statisticians on widely accessible computer files. Perhaps most of all, we need to acknowledge technical limitations, (e.g., the large degree of uncertainty that is often swept under the rug of unchallenged assumptions). If we duck these challenges, we will indeed be relegated to the role of purveyors of increasingly empty ritual.

CLOSING REMARKS

I have attempted to illustrate that probabilistic modeling, data analysis, and statistical inference can all contribute to exploratory research, given sufficient familiarization of the statistician with the scientific challenge. When the statistician is fortunate enough to become a member of the investigative team, the work will likely benefit from review by some other trained skeptic.

REFERENCES

Blank, Martin and Reba Goodman (1989). New and Missing Proteins in the Electromagnetic and Thermal Stimulation of Biosynthesis. Bioelectrochemistry and Bioenergetics: 11-21.

Box, George E. P. (1979). Some Problems of Statistics in Everyday Life. J. Amer. Statist. Assn. 74, 365: 1-4.

Goodman, Reba, and Ann S. Henderson (1988). Exposure of Salivary Gland Cells to Low-Frequency Electromagnetic Fields Alters Polypeptide Synthesis. Proc. Natl. Acad. Sci. USA 85: 3928-3932.

Goodman, Reba, C. Andrew Bassett, and Ann S. Henderson (1983). Pulsing Electromagnetic Fields Induce Cellular Transcription. Science 220: 1283-1285.

Lee, Casey K. (1989). Prospective Double-Blind Study of PEMFs in Interbody Lumbar Fusion. Abstracts from the 1989 annual meeting of the Bio-Electric and Magnetic Society (BEMS) Session C-1: Symposium on the Application of Time-Varying Fields for Tissue Healing.

Tukey, John W. (1986). Sunset Salvo. The American Statistician 40, 1: 72-76.

Tukey, John W. (1979). Comment on Nonparametric Statistical Data Modeling by Emanuel Parzen. J. Amer. Statist. Assn. 74, 365: 121-122.

Figure 1. Statistical activities (interior) supporting the cycle of scientific inquiry.

DISCUSSION Special Problems II

Richard L. Anderson
Statistical Consultants, Inc.
Lexington, Kentucky

STATISTICIANS AND EXPERIMENTERS

- Statisticians should have some applied scientific background, if possible.
- Statisticians should be involved at the start of an investigation.
- Statisticians should be careful to not be the principal investigators.
- Important contributions by statisticians include experimental design, selection of variables, estimation of number of observations needed for desired power, and evaluation of sources of variability.
- Data Analysis includes:
 - (i) Location and recommendations regarding outliers.
 - (ii) Are the data sufficient?
 - (iii) Data entry and computing procedures.
 - (iv) Comparison of models and tests of departures from assumed model.
 - (v) Write final report.
- Some Problems
 - (i) Conflicting outcomes: side effects and safety v. efficacy.
 - (ii) Mixed models: several variance sources, repeated measures and unequal error variances.
 - (iii) Use of plateaus (upper- and lower-threshold).
 - (iv) Use of *2-treatment 3-period crossovers instead of traditional 2-2 plan.
 - (v) Do we place too much emphasis on hypothesis tests?
 - (vi) Some SAS computer programs.
 - (vii) Interpretation of regression coefficients in multiple regression models.
 - (viii) Selection of *regressors or predictors.
 - (ix) Use of Fisher's exact test in analyses of 2 x 2 tables.

SOME COMMENTS ON MULTIPLE COMPARISONS

- History: Multiple Comparison procedures were constructed to compare selections from many treatments (n) with no defining characteristics. Duncan based his results on decreasing protection rates as the number of treatments to be compared increases:

$$P = (1 - \alpha)^{n-1}. \text{ If } \alpha = 0.05 \text{ and } n = 3, P = 0.9025; \text{ if } n = 10, P = 0.63$$

Then Duncan and Walker came up with a Bayesian procedure based on an F-test. Scheffe and Tukey (in particular) had the same protection rate regardless of n. It should be emphasized that these are all experiment-wise error rates. Since I feel that most experimenters have certain specific effects to compare (before the experiment is conducted), I usually prefer comparison-wise error rates, which do not consider the number of comparisons made.

- Multiple Comparison Procedures (see Carmer-Swanson, 1973; n treatments;
% = nominal error rate)

LSD = Least Significant Difference (comparison wise)
 = $t s(d)$, where $s(d)$ = standard error of mean difference.
 FSD = Fisher LSD (requires preliminary significant F-test).
 TSD = Tukey SD based on studentized range (experiment-wise)
 SSD = Scheffe' SD based on $(n-1)$ square root $[F_{\alpha}^2(d)]$ (experiment-wise).
 BET = Bayes' exact test (Waller-Duncan)
 MRT = Duncan multiple range test based on experiment-wise but decreased protection rates as number of treatments in comparison increases.

- Comparisons of % Correct Decisions for Various Procedures (Carmer-Swanson, 1973 Simulated Experiments)

$n = 10$ treatments; 5% nominal rate

Method	C(%)	E(%)	% correct Decisions (R^*, Z)		
			*3,2	*3,1	6,1
FSD	1.2	5.4	62.0	21.0	40.0
LSD	5.0	58.4	63.0	21.0	40.0
TSD	0.2	4.8	18.0	2.6	6.6
SSD	0.01	0.3	4.3	0.3	0.8
BET	1.5	18.4	63.0	22.0	45.0
MRT	2.6	37.3	56.0	17.0	34.0

Error rates (C: Comparison-wise; E: Experiment-wise).

R replications and $Z = \frac{\text{average treatment difference}}{\text{standard deviation}}$.

Note the high E for LSD, BET, and MRT and the low % correct decisions for TSD and SSD.

- Some Examples

(i) Total dry matter yield of 11 alfalfa cultivars (Petersen, 1977): multiple comparison procedure.

Cultivar	A	B	C	D	E	F	G	H	I	J	K
Yield	11.60	11.32	11.27	10.79	10.32	10.06	10.04	9.26	8.54	8.52	8.27
MRT*	a	a	a	ab	ab	abc	abc	bcd	cd	cd	d

(*Values followed by same letter are not significantly different at $\alpha = 0.05$.)

(ii) Total dry matter of stems as affected by plant age (Petersen, 1977): regression analysis, not multiple comparison.

Age (weeks)	4	6	8	10	12	14	16	18
Y = Dry Matter	24	28	35	48	44	49	51	51
(% Total)								
MRT	f	e	d	b	c	b	a	a
X_1	4	6	8	10	10	10	10	10
X_2	0	0	0	0	2	4	6	6

Used a linear-plateau procedure for which

$$\hat{Y} = \begin{cases} 7.38 + 3/71 X_1 & \text{for Age} \leq 10 \\ 44.49 + 1.034 X & \text{for } 10 \leq \text{Age} \leq 16 \\ 50.69 & \text{for Age} \geq 16. \end{cases}$$

- (iii) Simulated 2 x 2 Factorial plus check (Nelson and Rawlings (1983): not multiple comparison.

		<u>ANOVA</u>	
	Degrees of Freedom	Mean Square	F
Blocks		3	1900
Treatments			
ck v. others	1	2225	7.5*
Varieties (V)	1	1500	5.0*
Nitrogen (N)	1	1250	4.2
VN	1	825	2.8
Error	12	300	

(* $\alpha < .05$)

Note that pertinent comparisons, each with 1 d.f., are tested.

REFERENCES

- Carmer and Swanson (1973). JASA 68: 66-74.
 Chew (1976). Hort. Science 11*(4): 348-357.
 Duncan (1955). Biometrics 11: 1-42.
 Duncan (1965). Technometrics 7: 171-222.
 Dunnett (1955). JASA 50: 1096-1121.
 Nelson and Rawlings (1983). Jour. Agron. Educ. 12: 100-105.
 O'Neill and Wetherill (1971). JRSS(B) 36: 218-250.
 Petersen (1977). Agron. Jour. 69: 205-208.
 Scheffe' (1953). Biometrika 40: 87-104.
 Tukey (1953). Trans N.Y. Acad. Sci.; Series II, 16: 88-97.
 Waller and Duncan (1969). JASA 64: 1484-1503
 Waller and Duncan (1972). JASA 67: 243-5.

META-ANALYSIS

Kay T. Kimball
Baylor College of Medicine
Houston, Texas

Interest in biological effects of electric fields has motivated research in a variety of areas, but summary research statements have been complicated by the diversity of endpoints used in the studies. Often, the studies in common areas have reported inconsistent and conflicting results, consequently most of the critical reviews have been unable to identify consistent electric field effects. Most of the reviews have been based on qualitative evaluation of the studies, often supported by a "vote count," that is, counting the number of observed results in each of three categories: negative effect, positive effect, and nonsignificant result. When vote counts are used to review studies with small effect sizes, the power to detect a true effect tends to zero as the number of studies increases, even when sample sizes are large (1). When studies have small sample sizes (and consequently low power to detect small treatment effects), a vote count can be completely misleading, because the individual experiments are likely to have insignificant results even when real effects exist (a false negative result).

Meta-analysis, a statistical method for quantitative synthesis of a body of literature, was introduced in the social sciences in the 1970's and has recently been used with increasing frequency in the medical literature, in particular, to pool results of randomized controlled trials (2-19). Meta-analysis formally combines the results from studies that address the same question by pooling information over the studies. Better estimates of treatment effect can be obtained than those from the individual studies, especially when the individual studies lack the statistical power to detect a true treatment effect. Meta-analysis can also use the variability among studies to explore the size and direction of effects among exposure or age groups or different populations.

A meta-analysis consists of several stages that are common to the protocol for any research study: formulating a clear question, acquiring studies and applying selection criteria, abstracting data (usually summary statistics) and representing endpoints in a common metric, and statistical analysis and interpretation of results. Just as in primary research, a precise answer will not be obtained from an experiment based on a broad question. Naylor (20) suggested that if one wants to blend "apples" and "oranges" in an overview, information about "fruit" will be obtained. Sometimes, this may be the desired goal. For example, a broad question would be appropriate in cases where broad policy issues are being addressed. It is important to obtain an unbiased collection of studies that have addressed the specific question of interest. Because of the publication bias in favor of statistically significant results, a diligent search of published and unpublished studies is necessary. This is usually accomplished by computer searches, by consulting reviews, indices and references cited in articles and reports, and by communication with researchers in the field of interest. As studies are acquired, it is helpful to communicate with each investigator for suggestions about other studies and the availability of unpublished data. Unless all papers addressing the specified question are accepted, the selection process for studies to be included in a meta-analysis can strongly influence the outcome. Four meta-analyses on steroid use and peptic ulcers, which were carried out on the same body of literature, reported results varying from highly significant to totally nonsignificant; the only difference lay in their acceptance and rejection of papers (21). Chalmers (22) suggested a method of assessing the quality of a randomized controlled trial whereby the methods and results sections are disguised and assessed separately by a panel of judges.

The measures of effect size calculated for each study depends upon the type of data

that are being summarized. Studies that report means, standard deviations, and sample sizes will usually be combined using standardized mean differences. In meta-analyses of clinical trials and other epidemiological studies, risk ratios (13), risk differences (3), and odds ratios (10) have been pooled. The estimate from each study is often assigned a weight that is proportional to the precision of the study's results (usually the inverse of the variance of the risk estimate). Other measures of effect that have been used are differences between correlation coefficients (23) and probit transformations of proportionate data (24).

Meta-analysis of a collection of studies is similar to primary analysis of a collection of data points, and many familiar statistical techniques have been adapted for this application. When a group of studies can reasonably be expected to have a common effect size for an outcome (i.e., the sample effect sizes in the individual studies all estimate the true population effect size), the estimates of sample effect size can be pooled across studies to draw general conclusions. Tests of homogeneity should be applied to determine whether the studies share a common effect size or whether subsets of the studies have common effect sizes. Homogeneous subsets of effect sizes can be quantitatively summarized and statistical confidence statements obtained. The subsets should be determined a priori by characteristics of the studies. The effect sizes of the homogeneous subsets can also be compared in a manner analogous to analysis of variance and tests for significant differences among groups of studies can be made. Conversely, effect sizes can be examined directly for disjoint and overlapping groups. Common effect sizes for these clusters can be estimated, but should only be used to generate hypotheses for future studies about cluster characteristics. When the effect sizes are significantly heterogeneous, weighted regression analysis can be used to identify the factors that influence the variability among studies (25). For example, dose-response relationships can be evaluated. Hence, variety in study designs can be valuable in assessing why study results differ.

EXAMPLES OF QUANTITATIVE REVIEW OF ELF AND EMF LITERATURE

Two types of meta-analyses were used to review the results of 40 experiments on hematologic responses in laboratory animals exposed to electric fields. First, a meta-analysis was done that combined observed significance levels over the experiments for each endpoint. This type of meta-analysis is based on the assumption that when no true treatment effect exists, the significance levels will be uniformly distributed between zero and one. The results strongly suggested specific and consistent hematologic responses to electric field exposure on some endpoints examined and the need for further analyses of these data (26).

The second meta-analysis estimated the effect size for each endpoint in each experiment using the standardized mean difference between the control and the experimental group. Because the effect sizes for each endpoint varied considerably, the relationship between the effect sizes and characteristics of the studies was examined using a weighted least squares regression model with each effect size weighted by the inverse of its variance. Exposure intensities in these experiments varied between 0.2 and 240 kV/m, while exposure times varied between 15 and 250 days. Other characteristics included in the regression were investigator, type of animal, sex, and field frequency. In the regression models, total protein and eosinophils were significantly associated with levels of exposure to electric fields and two other endpoints yielded borderline significance levels (27).

Since most statistical procedures require independent observations, correlated observations can present problems. To avoid inclusion of correlated effect sizes in these meta-analyses, experiments that shared common control groups or which repeated measurements on animals used in other experiments were discarded. A separate meta-analysis was done for each endpoint, because hematologic endpoints within an animal would be expected to be correlated. Multivariate meta-analytic techniques have

recently been developed that adjust for correlations among effect sizes caused by inclusion of a common control group and multiple observations on the same subject (28). These new techniques will allow simultaneous testing of sets of hematologic responses, which are correlated instead of analyzing each endpoint separately, thereby giving better control over the type I error rate, and to include the experiments that were omitted from the earlier meta-analyses because of common control groups.

Savitz and Calle (29) reported a quantitative review on adult leukemia incidence and mortality in persons with occupational exposures to electromagnetic fields. They examined eleven studies (30-40) that had primarily reported proportional morbidity or mortality measures. All risk estimates were converted to observed and expected numbers of cases; these were summed across studies or across occupational categories to produce aggregated risk estimates that were weighted by the size of each population. The summary risk estimates showed consistently elevated risk ratios for selected occupations potentially exposed electromagnetic fields.

Coleman and Beral (41) combined the results of eleven studies (35-40, 42-46) on leukemia incidence and mortality in electrical workers. Using a procedure similar to Savitz and Calle, they calculated summary risk estimates from studies that primarily had reported standardized and proportional mortality and morbidity ratios. They found that electrical workers were at slightly increased risk of leukemia, with stronger associations for some types of leukemia. As noted by the authors, both of these reviews as well as the individual studies were inherently limited regarding the effect of EMF due to the absence of exposure characterization. Because direct measurements of exposure are not usually available, several methods have been suggested for classifying exposures based on job titles (42, 44-46). A more direct scheme for classifying these exposures may be available from two occupational hygiene studies currently in progress that are directly measuring EMF exposures in occupational groups in the field (47,48). If their results can be used to classify or rank exposures in the studies that now contain only job title information, a meta-analysis of these occupational studies could include exposure classifications.

META-ANALYSIS AS AN AID IN PLANNING STUDIES

Meta-analysis need not always be used retrospectively. A series of studies can be designed for later meta-analysis, for example, a program of studies can include several different study designs. The diversity of populations and exposures from several adequately sized studies can yield results that are more generalizable than those from one large study. Investigators in a series of studies carried out over time can learn from mistakes in earlier studies and can take advantage of new technology. Interim meta-analyses after some of the studies are complete can also provide useful information for planning the remaining studies.

REFERENCES

1. Hedges LV, Olkin I. Vote-counting methods in research synthesis. Psychological Bulletin 1980;88:359-369.
2. Colton T, Freedman LS, Johnson AL, eds. Proceedings of the Workshop on Methodologic Issues in Overviews of Randomized Clinical Trials, May 1986. Statistics in Medicine 1987;6(3).
3. DerSimonian R, Laird R. Meta-analysis in clinical trials. Contr. Clin. Trials 1986;7:177-188.
4. Gerbarg ZB, Horwitz RI. Resolving conflicting clinical trials: guidelines for meta-analysis. J. Clin. Epid. 1988;41:503-509.
5. Greenland S. Quantitative methods in the review of epidemiologic literature. Epid. Rev. 1987;9:1-30.
6. Louis TA, Fineberg HV, Mosteller F. Findings for public health from meta-analyses.

- Ann. Rev. Public Health 1985;6:1-20.
7. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. N. Engl. J. Med. 1987;316:450-455.
 8. Beaumont JJ, Breslow NE. Power considerations in epidemiologic studies of vinyl chloride workers. Am. J. Epid. 1981;114:725-734.
 9. Buyse M, Zeleniuch-Jacquotte A, Chalmers TC. Adjuvant therapy of colorectal cancer. Why we still don't know. JAMA 1988;259:3571-3578.
 10. Canner PL. Aspirin in coronary heart disease. Isr. J. Med. Sci. 1983;19:413-423.
 11. Collins R, Yusuf S, Peto R. Overview of randomised trials of diuretics in pregnancy. Brit. Med. J. 1985;290:17-23.
 12. Collins R, Langman M. Treatment with histamine H2 antagonists in acute upper gastrointestinal hemorrhage. N. Engl. J. Med. 1985;313:660-6.
 13. Stampfer MJ, Goldhaber SZ, Yusuf S et al. Effect of intravenous streptokinase on acute myocardial infarction. Pooled results from randomized trials. N. Engl. J. Med. 1982;307:1180-1182.
 14. Steinberg EP, Shepard DS. Lead—is it carcinogenic? Public Health Rev. 1983;11:177-192.
 15. Stewart AL, Reynolds EOR, Lipscomb AP. Outcome for infants of very low birthweight: survey of world literature. Lancet 1981;1:1038-1041.
 16. Tran ZV, Weltman A. Differential effects of exercise on serum lipid and lipoprotein levels seen with changes in body weight: a meta-analysis. JAMA 1985;254:919-924.
 17. Tyroler HA. Total serum cholesterol and ischemic heart disease risk in clinical trials and observational studies. Am. J. Prev. Med. 1985;1:18-24.
 18. Yusuf S, Collins R, Peto R, et al. Intravenous and intracoronary fibrinolytic therapy in acute myocardial infarction: overview of results on mortality, reinfarction and side-effects from 33 randomized controlled trials. Eur. Heart J. 1985;6:556-585.
 19. Yusuf S, Peto R, Lewis J, et al. Beta blockade during and after myocardial infarction: an overview of the randomized trials. Prog. Cardiovasc Dis. 1985;27:335-371.
 20. Naylor, CD. Two cheers for meta-analysis: problems and opportunities in aggregating results of clinical trials. Can. Med. Assoc. J. 1988;138:891-895.
 21. Chalmers TC. Meta-analysis of clinical medicine. Transactions of the American Clinical and Climatological Assoc, Ponte Vedra Beach, FL, 1987;99:144-150.
 22. Chalmers TC, Smith H, Blackburn B, et al. A method for assessing the quality of a randomized control trial. Controlled Clin Trials 1981;2:31.
 23. Cohen J. Statistical power analysis for the behavioral sciences. Rev ed. New York, Academic Press, 1977.
 24. Glass GV, McGaw B, Smith ML. Meta-analysis in social research. Beverly Hill, Sage Publications, 1981.
 25. Hedges LV, Olkin I. Statistical methods for meta-analysis. Academic Press, New York, 1985.
 26. Morris MD, Kimball KT, Aldrich TE, Easterly CE. Statistical approach to combining the results of similar experiments, with application to the hematologic effects of extremely-low-frequency electric field exposures. Bioelectromagnetics 1989;10:23-34.
 27. Kimball KT, Easterly CE. A meta-analysis of the hematologic effects of ELF electric field exposures on small laboratory animals. DOE/EPRI International Review of Research on 60 Hz Electric and Magnetic Fields. Kansas City, MO, 1987.
 28. Kimball KT. Multivariate methods for correlated effect sizes. Doctoral Dissertation, University of Texas School of Public Health, Houston, 1988.
 29. Savitz DA, Calle EE. Leukemia and occupational exposure to electromagnetic fields: review of epidemiologic surveys. J. Occup. Med. 1987;29:47-51.
 30. Dubrow R, Wegman DH. Occupational characteristics of cancer victims in Massachusetts 1971-1973. DHHS (NOISH) publication no. 84-109, Cincinnati, OH, National Institute for Occupational Safety and Health, 1984.
 31. Guralnick L. Mortality by occupation and cause of death among men 20-64 years of age: United States, 1950. Vital Statistics: Special Reports vol 53, no. 3, Government Printing Office, 1963.

32. Roth HD. An evaluation of published studies analyzing the association of carcinogenesis with exposure to magnetic fields. EPRI EA-3904, RP 1824-03. Palo Alto, CA, Electric Power Research Institute, 1985.
33. Peterson GR, Milham S Jr. Occupational mortality in the state of California 1959-61. DHEW (NIOSH) publication no. 80-104, Cincinnati, OH, National Institute for Occupational Safety and Health, 1980.
34. Milham S Jr. Silent keys: leukaemia mortality in amateur radio operators. Letter. Lancet 1985;1:812.
35. Milham S Jr. Mortality from leukemia in workers exposed to electrical and magnetic fields. Letter. N. Engl. J. Med. 1982;307:249.
36. Wright WE, Peters JM, Mack TM. Leukaemia in workers exposed to electrical and magnetic fields. Letter. Lancet 1982;2:1160-1161.
37. McDowall ME. Leukaemia mortality in electrical workers in England and Wales. Letter. Lancet 1983;1:246.
38. Coleman M, Bell J, Skeet R. Leukaemia incidence in electrical workers. Letter. Lancet 1983;1:982-983.
39. Pearce NE, Sheppard RA, Howard JK, et al. Leukaemia in electrical workers in New Zealand. Letter. Lancet 1985;1:811-812.
40. Calle EE, Savitz DA. Leukemia of occupational groups with presumed exposure to electrical and magnetic fields. N. Eng. J. Med. 1985;313:1477-1478.
41. Coleman M, Beral V. A review of epidemiological studies of the health effects of living near or working with electricity generation and transmission equipment. Int. J. Epid. 1988;17:1-13.
42. Howe GR, Lindsay JP. A follow-up study of a ten-percent sample of the Canadian Labour Force. I. Cancer mortality in males, 1965-73. J. Natl. Cancer Inst. 1983;70:37-44.
43. Olin R, Vagero D, Ahlbom A. Mortality experience of electrical engineers. Brit. J. Ind. Med. 1985;42:211-212.
44. Stern FB, Waxweiler RJ, Beaumont JJ, et al. A case-control study of leukemia at a naval nuclear shipyard. Am. J. Epid. 1986;123:980-992.
45. Tornqvist S, Norell S, Ahlbom A, Knave B. Cancer in the electric power industry. Brit. J. Ind. Med. 1986;43:212-213.
46. Wilklund K, Einhorn J, Eklund G. An application of the Swedish Cancer-Environment registry. Leukaemia among telephone operators at the Telecommunications Administration in Sweden. Int. J. Epid. 1981;10:373-376.
42. Lin RS, Dischinger PC, Conde J, Farrell KP. Occupational exposure to electromagnetic fields and the occurrence of brain tumors. J. Occup. Med. 1985; 27:413-419.
44. Wertheimer N, Leeper E. Electrical wiring configurations and childhood cancer. Am. J. Epid. 1979;109:273-284.
45. Spitz MR, Johnson CC. Neuroblastoma and paternal occupation. A case-control analysis. Am. J. Epid. 1985;121:924-929.
46. Hsieh CC, Walker AM, Hoar SK. Grouping occupations according to carcinogenic potential: occupation clusters from an exposure linkage system. Am. J. Epid. 1983;117:575-589.
47. Bowman JD, Garabrant DH, Peters JM. Assessment of occupational exposures to electric and magnetic fields. Presented at DOE/EPRI Contractors Review, Biological Effects from Electric and Magnetic Fields. Kansas City, MO, 1987. (Completion expected 1989)
48. Knave B, Floderus-Myrhed B, Tornqvist S, et al. Occupational hygiene study of workers exposed to low-frequency magnetic fields. Presented at DOE/EPRI Contractors Review, Biological Effects from Electric and Magnetic Fields. Kansas City, MO, 1987.

RESEARCH ON THE HEALTH EFFECTS OF ELECTROMAGNETIC FIELDS: SCIENCE, UNCERTAINTY, AND STOPPING RULES

John C. Bailar

Office of Disease Prevention and Health Promotion
U.S. Department of Health and Human Services
Washington, D.C.

and

McGill University School of Medicine
Montreal, Canada

For several years I have been an interested, sometimes amused, even sometimes bemused spectator at the unfolding of new findings from research on the health effects of electromagnetic fields. However, I am in no way an expert on electromagnetic fields or their biologic effects, but rather a research methodologist with a special interest in how people really learn from data. Perhaps this distance in subject matter, and perhaps some difference in perspective, will be helpful in assessing where we stand and where we should be going. At the end, I will give special attention to what I call "stopping rules"; that is, more or less formal determinations, generally before the fact, about what research outcomes will cause us to abandon a particular line of inquiry. It seems intuitively clear that consistently unproductive fields of study should not be studied indefinitely, but what do we mean by "unproductive," and how should we apply that concept to broader or smaller areas of investigation?

My observations of research related to the effects of electromagnetic radiation on human health lead me to four conclusions. First, standards of research in this field need to be substantially improved. Second, there is a need for a broad conceptual model about how to draw inferences regarding cause-and-effect relationships between electromagnetic exposures and human health. Third, there is a need for a few "anchors"—solid demonstrations that certain effects do or do not exist at low exposure levels; this will depend on having an appropriate conceptual model. Finally, we should develop a rational way to set priorities for future work; thus will depend on both having a conceptual model and developing several anchors. The need for stopping rules comes directly out of the need for better ways to set priorities. The remainder of my talk will develop these four conclusions.

STANDARDS OF QUANTITATIVE ANALYSIS: STATISTICS AND EPIDEMIOLOGY

Statistics and its daughter discipline epidemiology are critically important in many kinds of research on human health, and I believe that they are so here. Debates over the boundaries of one or another discipline are generally unproductive, and I offer my own choice of definitions only to help readers understand what I mean when I use the words:

Statistics—the art and science of interpreting quantitative data (including counts of events that are subject to uncertainty).

Epidemiology—the branch of applied statistics that studies the determinants and correlates of human disease.

In the opening talk here, reference was made to the need to define the role of the statistician, and by implication the epidemiologist, in studies of the health effects of electromagnetic radiation. This role includes such diverse matters as pursuing the specific questions for study, drawing up protocols, maintaining data quality, reducing results to a manageable size, performing various kinds of "statistical" computations, and generalizing results from the study group to broader populations. Thrall asked, "How much can statistics help?" A lot, if one adopts this broad definition. For many decades,

epidemiology has been a community health search on insight into the causes, and in many

cases the prevention, of human illness and death. Some examples of epidemiologic links of various strength between exposure to some agent and health outcomes are:

Conclusive:	Cigarettes and lung cancer, some industrial carcinogens (asbestos, vinyl chloride, etc.) and cancer; fetal alcohol syndrome
Good:	Diet (cholesterol, fat) and heart disease; highway speed limits and accidental death/injury
Fair:	Dietary sodium and hypertension; diet and cancer; altitude and reproductive problems
Poor:	Exposure to electromagnetic fields

As a scientific discipline, epidemiology (like statistics more generally) has the unfortunate appearance of being easy. Not so. It is true that both exposures and outcomes can generally be put in simple terms, but there are almost always serious problems in such matters as selection of appropriate subjects and controls, minimizing a wide range of possible biases in the data, and drawing cause-and-effect inferences from strictly observational (non-experimental) data. Relevant concepts, techniques, and skills can be taught and developed, and trained epidemiologists or biostatisticians can generally add a great deal of strength to epidemiologic inquiries. Unfortunately, much past work on electromagnetic fields and human health has not met state-of-the-art standards; this may be one important reason why so little has been settled.

CONCEPTUAL MODELS

Epidemiologic inquiry can answer some questions about the effects of noxious exposures on human health, but not others. For many years a useful rule of thumb was that a cause-and-effect relative risk of 2 or greater (that is, a doubling of risk in an exposed population over similar persons not exposed) could often be detected and demonstrated by epidemiologic data, but that smaller relative risks might well be the result of covariates ("confounders"), recognized or unrecognized. For example, the link between cigarette smoking and lung cancer (relative risk of 10 or more) was accepted as cause-and-effect many years before the similar link between cigarette smoking and cardiovascular disease (relative risk on the order of 2), despite the fact that doubling a common condition such as cardiovascular disease is a far greater health burden than a 10-fold increase in a much less common disease like lung cancer.

Recent advances in epidemiologic concepts and methods have reduced this lower bound on interpretable relative risks from 2.0 to perhaps 1.5, but only for large, well executed studies, that have been designed by very capable professionals and with special attention to the full range of possible biases and other problems. What are the characteristics of epidemiologic data that hamper the conclusive interpretation of relative risks smaller than, say, 2.0 for the ordinary (but state-of-the-art) epidemiologic study? First, as already noted, epidemiology is inherently a descriptive rather than an experimental science. There are rarely opportunities for an investigator to manipulate exposures of human populations "in the wild"—that is, under conditions other than special short-term clinical or laboratory settings, where conditions may differ from the usual milieu.

A second, related difficulty is that study populations must often be quite large; this includes case-control studies if "population" is defined to mean the source population of both cases and controls. And this in turn causes problems of control over data quality as well as (sometimes) astonishingly high costs. Because of costs and other problems, it is understandable that epidemiologic studies are sometimes designed to answer several different kinds of questions, with consequent compromises in design. Sometimes the limitations of sample size simply cannot be resolved regardless of resources available, because study populations or affected persons are not sufficiently numerous, or because needed historical information is not available. For example, research on the health of

humans exposed to formaldehyde has been seriously hampered by the small numbers of heavily exposed persons, the lack of accurate information about their specific levels of exposure to formaldehyde, and possible concurrent exposures to other chemicals.

In designing a research attack on any given question about electromagnetic fields and human health, one may consider several possible epidemiologic studies as well as a range of animal studies of varying relevance. Often, more than one study is conducted, especially if results are seen as important. Thus we may have more than one set of results, and the results may not agree. There have been two general approaches to dealing with multiple (and often conflicting) sources of data. These correspond roughly to the concept of a "definitive study" vs. meta-analysis. One has been to screen each report against some set of criteria of quality, and to use only the one study (or the very few studies) that seem strongest. An example is the widespread practice, in the U.S. Federal Government and elsewhere, of basing quantitative estimates of health risks on the single report, human or animal, that is judged to be best for the specific purpose. I do not criticize this approach because I recognize the grave problems of estimating a single dose-response relationship from studies conducted in dissimilar ways and yielding quantitatively incompatible results. In general, however, and as long as the focus of inquiry is on whether there is any risk at all, I prefer an alternative approach of using the whole of the evidence, including subjective weightings for reliability of findings and conclusions. Opinions may often differ about how to assess individual studies and about how to interpret the whole of the evidence, but it is generally inappropriate to try to find a central study or core of studies that largely agree, and to just set aside any others that seem incompatible. Quite the opposite, one should generally pay special attention to the outliers—those that are at variance with the main body of findings—until each incompatible finding has been adequately explained. I regret the need to add a caution that the study of outliers must be strictly objective, a standard not always met in epidemiology as a whole, or in studies of electromagnetic exposures and health.

A fourth broad point is that the evidence as a whole bearing on "big" questions of health almost always has four prominent features: It is vast; it is complex; it is largely to one side of the central questions (e.g., we use animal studies in lieu of human experimentation); and it is mostly of poor quality. This is true, for example, of the evidence relating human health to diet, to air pollution, and to electromagnetic fields. These features of the data have been discussed elsewhere (1) and I will not expand here.

These numerous and serious difficulties point to the need for a conceptual model—that is, a theoretical framework for use in sorting and organizing what is relevant, understanding which problems are serious and which are not, and more generally comprehending all that can be brought to bear on inferences and conclusions. A conceptual model of this sort can be an invaluable aid in attacking very complex problems, but developing a model can require much time and effort, as well as a detailed understanding of the full range of relevant subject matter. Thus, high-level leadership is generally needed, and to be effective, a model usually needs a broad support from all of the interested parties. I urge these parties to cooperate in developing a conceptual model for possible health effects of electromagnetic fields.

Two other conceptual models may be helpful as illustrations of what I mean by a "conceptual model" and as possible sources of elements for a new model for the present topic.

First is the model developed by a committee of the National Academy of Sciences (2) as a general framework for quantitative health risk assessment. This model has undergone some refinements around its edges (e.g., 3), but the central idea remains intact: There are four largely separable steps, each of which requires separate development. These are listed in Table 1. A particular value of this model is that it has broken the very complex problem of risk assessment into four pieces, each critical to the outcome but each also more open to study by an appropriate group of analysts. This

4-step approach has been widely adopted in the U.S. and elsewhere, with a broad consensus that it tends to make risk assessment simultaneously easier in a technical sense, more accurate (despite remaining major uncertainties), and more defensible against attack from the right or left.

A second example of a conceptual model is the set of empirical criteria developed by Hill (4) for drawing causal inferences from observational data. Again, the model has undergone some development: Table 2 shows the six criteria (derived from Hill) that were used in a recent report on relations between diet and health (5). Several other speakers here have already referred to this. No one element of this model is essential to a specific causal inference, but some two decades of experience with a wide range of substantive problems has shown that the model is a broadly useful way to organize the collection, presentation, and interpretation of observational studies on environmental exposures and health.

Both of these examples (Tables 1 and 2) have direct applications to questions about electromagnetic fields and health, but they are not entirely suitable, separately or together. For example, a special, additional problem not adequately addressed in either model is the vast range of possibly relevant exposure measures as well as possible outcomes, and the consequent problems of interpreting multiple comparisons. There are electric fields, magnetic fields, and associated charged particles and air ions; there are DC lines and AC lines, and the latter come in a full range of frequencies; and there are short-term and long-term exposures, with everything between, including continuous vs. variable exposure and even "windows." Several others here have already talked about the great problems that still remain in deciding just what to measure—or more specifically, what to use in data analysis (Bracker, Kheifitz, and others). Unfortunately, we do not seem to know enough yet to answer even the simplest of these questions, so for some time we will need to continue relying on crude epidemiologic descriptors. We cannot yet even follow the suggestion that we focus on dismissing some possible exposure metrics.

In addition to this multiplicity of possibly relevant aspects of exposure are the numerous possible outcomes: cancer, birth defects, respiratory problems, psychologic effects, and others, each with a plethora of more specific health endpoints. With all these possibilities it is virtually certain that some chance associations will be found to meet the usual standards of statistical significance ($p < 0.05$) even when there is in fact no effect. Further, it is almost as certain that in some instances, follow-up studies to confirm or refute earlier findings will themselves, just by chance, fall below the magic 5% value. This problem of multiple possible comparisons seems to me to be a major barrier to further understanding of electromagnetic fields and health, and should probably be addressed in any conceptual model for use in this subject area.

Another special issue is the very sparse data base on human effects combined with serious questions about the relevance of some possible animal studies. For example, fields induced in a 170 cm human may be far different, and have different biologic effects, from those induced in a 2 cm rodent. Overall, there seem to be special problems in defining the roles of both human and animal data that are not addressed in either of my two examples of conceptual models.

A third critical area, which certainly does arise in applications of those models but is not addressed directly by either of them, is the uncertainty that is inherent in all studies of chemical or physical exposures and health effects. This goes very far beyond the uncertainty expressed in the usual sorts of p -values and confidence limits. Such uncertainty is large—commonly several orders of magnitude, and sometimes many orders (6, 7)—and a good understanding of its sources and magnitude is clearly a part of the information critical to decision-makers who must use the scientific interpretations of data about one or another risk. However, we have no good way to estimate that uncertainty in specific instances, no good way to verify or validate our estimates, and no good way to convey our estimates of uncertainty to either non-scientific decision-makers or the public

at large. Issues of uncertainty might well be addressed in a conceptual model of electromagnetic exposures and human health.

Several previous speakers here (Sahl, Theriault, Thrall, and others) have already contributed other elements of a conceptual model for studies of electromagnetic exposure and health, and their comments reinforce my general feeling that a satisfactory, useful model could be developed.

ANCHORS

We have seen the promulgation of a large number of hypotheses, of varying specificity and with widely varying support from data, about the effects of electromagnetic field on health. Unfortunately, it seems that none of these hypotheses have yet been conclusively established or conclusively refuted. This leaves us in a situation, technically unsatisfactory and even awkward, of having no clear-cut standards for evaluating and improving methods of study (do specific investigations of new research methods show what we think they should, and if not why not?) and no clear-cut standards for interpreting evidence (does the strength of evidence for or against effect X approach that bearing on some accepted effect Y?). This has a further non-scientific but serious effect of leaving the whole field of study in some disrepute—few good scientists want to wade into a morass where nothing is ever settled—and it can even be self-propagating. ("Ho, hum, another study that no one will believe"—so, of course, no one believes it.) There are indeed many barriers to establishing such anchors. Table.3 lists some of those that seem most serious.

Getting conclusive results in this field, pro or con, will not be easy, but I believe that a few such results are necessary as an opening to further, better-targeted, better-accepted research. It will be particularly helpful if anchors of opposing types can be established for closely related conditions, so that one can say with some degree of assurance that exposure causes effect X but not effect Y, or at least not as much of Y. Possible examples (presented without judgment of the data) are leukemia vs. lymphoma, spina bifida vs. microcephaly, and childhood asthma vs. asthma in adults. A clear demonstration that any one of these is related to electromagnetic exposure while the other condition is not related might do much to settle controversies about both research methods and research findings.

I recognize that some informed observers may believe that we already have a few anchors—that one or another effect at low exposures may already be established, or that others may be adequately refuted. I do not want to get into specifics, but rather to point out that "established" and "refuted" refer to states of mind—individual interpretations, though they may integrate the views of others—and that in this sense little or nothing is settled. Your views here may be entirely correct, but what matters in the context I discuss is the opinions, right or wrong, of your scientific and technical colleagues, public interest and industry groups, regulators and other decision-makers, and the general public. Anchors depend on a broad consensus that we do not yet have.

SETTING PRIORITIES

The fourth broad need in this field is for improved ways to set priorities for future work. In the face of competing demands for support to pursue an impossibly large number of tantalizing leads, there is a major risk that funds will be allocated in ways short of optimal (almost without regard to the difficult question of defining "optimal"), and that the total pot of resources for distribution will also not be optimal. It should be entirely understood that I am addressing this issue from the point of view of the major research sponsors, not from that of the individual investigator, who must remain free to pursue new scientific knowledge within the constraints of available resources and support. These views—from the sponsor and from the investigator—are not incompatible because I expect the institutional priorities to change in response to new knowledge, and to retain some

degree of flexibility and some seed money for unusually promising work under the guidelines of the day. In addition, it seems likely that the numerous smaller, generally private, supporters of research will not agree on priorities, and of course some findings relevant to EM exposures and health may come out of work on other topics that is (legitimately) supported in other ways. The need, then, is not for a set of binding yes-or-no decisions about what will be supported, but a way for the major players to assure that most of their support is targeted toward the questions where important progress is most likely. Such targeting is not yet visible, but would be quite in line with support for other kinds of health research. For example, NIH Institutes are funded at widely different levels (more for cancer than for aging), and within the Institutes there has been much use of mechanisms to channel support to topics judged to be underdeveloped (though still to the best proposals within those topics).

Here also I will not be specific in developing the details of a recommendation, but some of the issues I would consider important in developing priorities for future work are listed in Table 4.

To this I would add Thrall's point that we need special attention to questions of which studies to replicate, and with what variations. Many other questions could be added. However, a critical underlying problem is the need for "stopping rules." How can we tell when we have, collectively, gone far enough in the pursuit of some hint or suggestion of an effect? When should there be a general willingness to abandon, at least for the present, intensive work on some interesting matter with a limited likelihood of payoff? We simply cannot do everything at once; we need comprehensive, integrated, rational processes to decide what to do first.

Thus, I return at last to the point I started with: stopping rules. Perhaps the first step in rationalizing research on EM exposures and human health endpoints is to develop some general standards or criteria about what kinds of research findings will not justify further research. Such "rules" will almost certainly have to be prospective, phased in with the initiation of new work, or even at the start of the following "generation" of research proposals. It is not too soon, however, to begin developing a consensus among all of the stake-holders, including the research community, about which possible future findings should be taken to justify follow-up studies and which should discourage further work (unless, and until, something new and unexpected turns up). It is also time to start asking research investigators, at the time of a request for support, what the possible outcomes of the research may imply regarding new studies. If most of the conceivable outcomes will simply lead to more proposals along the same lines, there may be a serious problem. The phrase, "along the same lines," is crucial; a real theoretical breakthrough would of course generate a range of new studies that could hardly be foreseen. But very, very few research findings can be described as breakthroughs, and none are in hand or on the horizon in studies of EM exposures and health.

To recapitulate, I recommend four steps, each of which will require a broadly based effort by a range of interested parties: improvement in the technical and scientific standards of research on electromagnetic fields and health, a broad conceptual model to bring such studies into appropriate relations with each other, a few solid anchors, and a mechanism to develop priorities. These steps should provide a sound basis for further research progress in this field.

Implementation? I will not be specific because I am convinced that the full range of interests must work together to develop a mutually acceptable approach to these four points. There seems to me to be a need for a very broadly based working group, of perhaps 9 to 12 persons, to carry forward the planning for work that will grow out of what we have heard at this meeting and elsewhere. Such a working group must be broadly based in at least two dimensions: technical discipline (Bracker has rightly made a major point here), and employment and orientation (industry, government, academic, public interest). Some members should also be knowledgeable about such matters as risk assessment, public information and risk communication, and public policy.

REFERENCES

1. Bailer JC, Thomas SR. What are we doing when we think we are doing risk analysis? Basic Life Sciences 33:65-76, 1985.
2. National Research Council. Risk Assessment in the Federal Government: Managing the Process. National Academy Press, Washington, D.C., 1983.
3. McGinnis JM, Bailer JC, Tardiff RG, Eds. (Task Force for Health Risk Assessment). Determining Risks to Health: Federal Policy and Practice. Auburn House, Dover, MA, 1986.
4. Hill AB. Principles of Medical Statistics, 9th ed. Oxford University Press, New York, 1971.
5. National Research Council. Diet and Health. National Academy Press, Washington, 1989.
6. Needleman J, Berney BL, Bailer JC, McGinnis JM. Methodological Challenges in Health Risk Assessment. Office of Disease Prevention and Health Promotion, U.S. Department of Health and Human Services, Washington, 1988.
7. Anderson P. Scientific origins of incompatibility in risk assessment. Statistical Science 3:320-327, 1988.

TABLE 1
COMPONENT STEPS IN HEALTH RISK ASSESSMENT

NRC (1983)	TFHRA (1985)
Hazard identification: The determination of whether a particular chemical is or is not causally linked to particular health effects.	Hazard identification: The qualitative indication that a substance/condition may adversely affect human health.
Dose-response assessment: The determination of the relation between the magnitude of exposure and probability of occurrence of the health effects in question.	Hazard characterization: The qualitative and quantitative evaluation of the nature of the adverse effects, including their expression as functions of the amount of exposure (dose).
Exposure assessment: The determination of the extent of human exposure before or after application of regulatory controls.	Exposure characterization: The qualitative and quantitative evaluation of the degree of human exposure likely to occur.
Risk characterization: The description of the nature and often the magnitude of human risk, including attendant uncertainty.	Risk determination: The integration of these steps into a scientific determination of the level of risk as a basis for policy consideration.
Abbreviation: NRC = National Research Council (Ref. 2), TFHRA = Task Force on Health Risk Assessment (Ref. 3).	

TABLE 2
A CONCEPTUAL MODEL FOR INFERENCE OF CAUSE-EFFECT RELATIONSHIPS

Some criteria for assessing the strength of cause-effect inferences from observational data:

Overall strength of association
Dose-response relationship
Specificity of association
"Correct" time relationships
Consistency of findings
Biologic plausibility

TABLE 3
SOME BARRIERS TO ESTABLISHING ANCHORS

Multiple and largely ill-defined hypotheses
Sufficiently large samples of heavily exposed humans
Possible "windows"
Uncertainty about appropriate dose metrics
Poor quality of most available work (not always the investigator's fault)
Highly multidisciplinary nature of needed work

TABLE 4
SOME POINTS FOR USE IN DEVELOPING PRIORITIES

How good is the present evidence that there is really something to look for—that there may really be some specific adverse effect of electromagnetic exposure on health?
How large is that effect likely to be (dose-response, including possible threshold)?
How many people may be affected?
How likely is it that important new advances in knowledge are within reach?
Who is affected (including possible synergies)?
Are victims individually identifiable?
Are results likely to apply to a single exposure-effect linkage, or more broadly?

BAYES METHODS FOR COMBINING DOSE-RESPONSE STUDIES

William DuMouchel
BBN Software Products Corporation
Cambridge, Massachusetts

Abstract

This paper discusses Bayesian statistical methods for combining or summarizing the results from multiple dose-response studies. The problem is treated as one of metaanalysis. A common complication arises when the dose-response studies are based on different species, or when it is desired to extrapolate from animals to humans. The Bayesian modeling process starts out by specifying components of variation corresponding to a) expected differences between studies, b) unpredictable differences between studies, and c) sampling variability within each study. Next, the available scientific knowledge about these sources of variation is incorporated into the model by the use of prior distributions on the parameters of the model. Standard computations for a Bayesian hierarchical model then lead to the posterior distributions of the parameters which are used to answer questions of interest. After reviewing two dose-response metaanalyses from the literature, the paper concludes with a general discussion.

1. Metaanalysis

Metaanalysis refers to the compilation of a review or synthesis of several presumably related studies or experiments, performed by different researchers, to come to a single overall conclusion. Critical judgements about the accuracy, quality and/or methodologies of the individual studies are necessary in order to combine them intelligently.

Qualitative metaanalyses concentrate on reviewing a set of studies of a particular problem with emphasis on classifying them by qualitative characteristics such as the kind of measurements taken, the potential biases they were subject to, and the degree to which the conclusions drawn by the original authors were justified, as well as what those conclusions were. The goal of such a research review may be to come to a general conclusion as to which of several competing hypotheses seem most supported by the preponderance of the evidence in the literature.

The focus of a quantitative metaanalysis is on using the specific statistical information from a set of studies in an overall statistical analysis to provide quantitative summary results. For this purpose it may be necessary to conduct an informal, qualitative, metaanalysis of a larger set of studies first, in order to select that subset of studies deemed suitable for the quantitative metaanalysis. For example, one might only use studies thought to be free from serious and hard-to-quantify biases, and for which certain necessary statistical summaries are available. Because a quantitative metaanalysis concentrates on performing a single statistical analysis of data from all the selected studies at once, the term *superanalysis* might be a more appropriate term. The book by Light and Pillemer (1984) provides many examples and discussions of both qualitative and quantitative aspects of research reviews.

Classical vs Bayesian Metaanalysis. The book by Hedges and Olkin (1985) surveys classical statistical methods of metaanalysis. The articles by Louis, Fineberg and Mosteller (1985) and Sacks et al. (1987) each review many metaanalyses in the areas of public health and clinical trials. Most authors recommend that a review of the research literature be conducted much like any statistical survey: Define a protocol for inclusion into the review, document the search for studies and list those found and those excluded for various reasons, address questions of combinability and potential biases of the studies used, and define a recognized statistical methodology to be used to combine the results. Preliminary descriptive summaries and graphical displays comparing the studies' results should precede the formal statistical pooling methodology. Of particular concern is the so-called "file-drawer problem", which occurs if those studies which happen to show statistically significant effects have a much greater chance of being published than do studies without significant effects, which remain hidden in their investigators' file drawers.

This paper focuses on formal methods of quantitative synthesis, assuming that the important preliminary issues mentioned above, and discussed at length by the referenced authors, have been satisfactorily handled. Many classical statistical methods for combining studies have been used. They include merely counting the number of studies which find a significant result, computing an average effect size across all studies, and computing a single combined level of significance for a set of statistical hypothesis tests, one from each study. Unfortunately, methods for combining significance levels from different studies do not distinguish between small studies with large effects and large studies with small effects. An important distinction is whether the method of synthesis assumes that each study is measuring the same underlying parameter. A common recommendation is to first conduct a formal hypothesis test of between-study homogeneity; only if this test is not significant is one advised to estimate the supposedly common parameter uniting the studies. This recommendation ignores the distinction between failing to reject interstudy homogeneity and proving that it is present. Methods of synthesis which explicitly allow for a component of variance for study-to-study variation avoid this logical

fallacy. Such methods usually require more detailed information from the primary studies than do other metaanalytic techniques, but they also usually provide more useful results for scientific and policy purposes.

The hierarchical Bayesian approach is distinguished by the construction and use of a formal statistical model at two levels. First, a parametric model is set up for each of the individual studies, in which a likelihood function relates the distribution of the sample statistics to one or more unknown parameters characterizing that study. Next, a second parametric statistical model is constructed which relates the parameters from the separate studies to each other. The computations involved in combining the two levels of the model are derived by the application of Bayes' formula for conditional probability.

The Bayesian approach is also usually distinguished by having little or no emphasis on hypothesis testing, as opposed to estimation. The Bayesian superanalysis provides a posterior distribution for each of the study-specific parameters; this will be a satisfying goal if the models are defined in terms of meaningful parameters. Research questions must be phrased as questions about the values of these parameters. The Bayesian model allows data from some studies to assist in the estimation of parameters characterizing the other studies. This notion of different studies "borrowing strength" from each other is a crucial one. However, in most realistic problems the degree to which different studies are relevant to each other is somewhat uncertain. Hence all but the simplest hierarchical Bayesian models also use the data to help decide *how related* the different studies are. Depending on whether or not the estimates from each study accord with the prior beliefs reflected in the model relating the different studies, the studies may or may not be able to borrow much strength from each other.

This last fact leads to the conclusion that the model relating the study-specific parameters should embody as much knowledge as possible about the problem area. Knowledge about mechanisms known to relate the different parameters, as well as more subjective knowledge, should be used, as long as the uncertainty of such estimates can be fairly expressed and incorporated into the model. These somewhat abstract principles are illustrated in the examples, where it should be clear that the ability to incorporate such subjective knowledge into the analysis is a strength of the Bayesian methodology. See DuMouchel (1989) for a more complete exposition of Bayesian metaanalysis and more references.

2. Combining Studies in Humans and Other Species

If the dose-response studies to be combined differ in the biological system being studied, the challenge of combining them is greater, of course. The Bayesian metaanalysis which synthesizes all the studies requires careful construction, since it is unlikely that dose-response curves from different species will be exactly equal. But Bayesian hierarchical models provide a natural way of attacking this "mouse to man" problem.

To carry out a hierarchical Bayesian analysis on the data from many studies, the data available from each must have a certain communality or comparability. The requirement of a second-level model relating parameters from the different studies implies that the estimates of these parameters, based on data within each study, have related meanings. For example a series of carcinogenesis studies may each estimate a relative risk or a standardized mortality ratio (SMR). Even though all the different SMRs might not reasonably be hypothesized to be equal, certain relationships between them may be plausible enough to allow a reasonable statistical model for them to be constructed.

Each study may estimate the slope of a dose-response curve. Even though the slopes from different studies may be expressed in widely varying units, it may still be possible to hypothesize approximate relationships between the parameters, as shown in the next section.

In addition, the methodology requires that each of the estimates be accompanied by its (approximate) standard error.

Different variations of the model could accommodate diversity in the type of study or in the biological endpoint used to measure a response. Most importantly, the method does not require the assumption that the animal data is relevant to humans. Instead, the degree to which the animal and human data agree can be estimated from the analysis.

These models do not address the other major problem which bedevils dose-response risk assessment, namely the low-dose extrapolation problem. If all the studies to be combined feature very high dose levels, then determining the shape of the dose-response curve at low dose levels is a key problem which is not discussed further here.

3. Example

The paper by DuMouchel and Harris (1983) presents an integrated analysis of 37 dose-response studies. The purpose of the research was to use the nonhuman studies to help assess the lung cancer risk to humans from long-term exposure to diesel fumes. The table shows a classification of the studies by biological system and chemical substance. The five biological systems include lung cancer in man, skin tumor initiation in mice, enhancement of viral transformation in hamster embryo cells, and mutagenesis in mouse lymphoma cells, with and without the addition of a metabolic activator. The ten substances tested were primarily polyaromatic hydrocarbons taken from various environmental emissions. There were diesel emissions from five types of diesel engines, plus five other substances with more or less similar chemical makeups. See DuMouchel and Harris (1983) for more information on the studies used.

As indicated in the table, there was only one available epidemiologic study involving diesel emissions, and this one study was not very large. It was desired to improve the risk estimates from that study, and to produce human lung cancer dose-response estimates for the other diesel emissions, even though no human studies of them were available.

If y_{ij} is the estimated log slope from the study of the j^{th} substance on the i^{th} biological system, and c_{ij} is its standard error, then the model

$$\begin{aligned} y_{ij} | \theta_{ij} &\sim N(\theta_{ij}, c_{ij}^2) \\ \theta_{ij} | (\alpha_i, \gamma_j, \sigma) &\sim N(\alpha_i + \gamma_j, \sigma^2) \end{aligned} \quad (1)$$

can be used to estimate the θ_{ij} s, the true log slopes, where α_i is an average log slope for the i^{th} biological system, γ_j is an effect due to the j^{th} substance, and σ is a measure of how poorly the true log slopes fit this additive setup.

The figure helps show the rationale behind the analysis. Plotted on vertical logarithmic scales in the figure are the estimates of the dose response slopes (with error bars corresponding to ± 1 standard error) for just three of the biological systems and three of the environmental emissions. Since the three vertical axes are drawn to the same logarithmic scale, and since equation (1) represents the mean of each θ_{ij} as $\alpha_i + \gamma_j$, the mean values of the log slopes should lie on three parallel lines, as shown on the figure. It is clear from the fact that the deviations of the points from the lines are often much greater than the indicated sampling error, the height of the error bars about each point, that there is additional variation in the observed log slopes that cannot be accounted for by sampling error or by overall effects of biological system or environmental emission. The key objective is to estimate the proper uncertainty to attach to the extrapolation in the upper right portion of the graph, the dose-response slope when humans are exposed to Diesel A. In addition to the uncertainty in the placement of the dotted lines in the graph, corresponding to uncertainty in $\alpha_i + \gamma_j$, the deviations of the observed data points from the fitted lines in the figure suggest that there is approximately an additional factor of two of uncertainty due to the expected deviation of the human-diesel A point from the overall trend.

When the model of equation (1) is fit to all of the data, the value of σ is estimated to be 1.1, meaning that extrapolation of a slope is typically too high or too low by a factor of $e^\sigma = 3.0$, and sometimes by the factor $e^{2\sigma} = 9$. However, within this factor of uncertainty, useful extrapolation is possible.

One obvious feature of the set of environmental emissions making up the rows of the table is that five of the ten are various types of diesel. In order to use this knowledge in the analysis, a prior distribution on the values of $\gamma_1, \dots, \gamma_{10}$ was used which expressed belief that the values of $\gamma_6, \dots, \gamma_{10}$, which are the average effects of the five diesel emissions, are relatively close together. In the analysis of DuMouchel and Harris (1983), the prior distribution used satisfied the condition that there was about a 25% prior probability that the dose-response slopes of any two of the diesel emissions differed by as much as a factor of ten. Using this prior specification the resulting 95% intervals for the relative risk to humans of lung cancer, assuming a 20 year occupational exposure to either diesel A or diesel E emissions, are:

Diesel E
(based only on the study of diesel E) (1.007, 3.95)

Diesel E
(Bayesian posterior distribution) (1.003, 1.31)

Diesel A
(Bayesian posterior distribution) (1.001, 1.28)

There is a dramatic reduction in uncertainty due to use of the Bayesian model.

4. Discussion

We have reviewed the concepts underlying Bayesian metaanalysis and discussed its use in combining several dose-response studies. In the example, linear dose response models were used, but the technique could be applied to nonlinear models just as easily, so long as parallel estimates are available from each of the studies being combined. Sometimes the linear-effect parameter can be thought of as a descriptive summary of a study rather than coming from a mechanistic model. In such cases it may be of interest to quantify the degree to which this natural measure of dose-response potency is consistent across all studies, even in the absence of belief in a linear model.

Of course the method is only as good as the component studies being combined. Often there are deficiencies in the reporting or the statistical analyses of the component studies, so that the hierarchical Bayesian analysis cannot be carried out because the required statistical analysis was not performed by the original authors. This can necessitate the cooperation of the original researchers and the reanalysis of some of the component data sets before the metaanalysis can begin. The construction of the Bayesian model can be severely constrained by the types of data available from the studies being reviewed.

Because of the statistical nature of the technique, in the absence of strong prior information, many studies are needed in order to make useful interspecies extrapolations. Sometimes the lack of a critical study is important. For example, in the metaanalysis of DuMouchel and Groër (1988), which combined 15 studies of the effects of isotopes of radium and plutonium on bone cancer, the lack of a rat study using radium severely affected the usefulness, for extrapolation to humans, of the rat studies using plutonium. We need to think more in terms of ensembles of studies and how each new study fits into the pattern of previous studies.

References

- DuMouchel, W. (1989) Bayesian Metaanalysis. in *Statistical Methods for Pharmacology*, Donald Berry, ed., New York: Marcel Dekker.
- DuMouchel, William and Groër, Peter (1988) A Bayesian methodology for combining radiation studies. *Proceedings of the Conference on Modeling for Scaling to Man*, Richland, WA: Battelle Memorial Institute. (to be published in *Health Physics*, 1989. A version of this paper appears as a chapter in the N.A.S. report *Biological Effects of Ionizing Radiation IV-α*, 1988.)
- DuMouchel, William and Harris, Jeffrey (1983) Bayes methods for combining the results of cancer studies in humans and other species (with discussion). *Journal of the American Statistical Association*, 78: 293-315.
- Hedges, J.V., and Olkin, I. (1985) *Statistical Methods for Metaanalysis*. Orlando, FL: Academic Press
- Light, R.J. and Pillemer, D.B. (1984) *Summing Up: the Science of Reviewing Research*. Cambridge MA: Harvard University Press
- Louis, T.A., Fineberg, H.V., and Mosteller, F. (1985) Findings for public health from metaanalyses. *Ann. Rev. Public Health* 6:1-20.
- Sacks, H.S., Berrier, J., Reitman, D., Ancona-Berk, V.A., and Chalmers, T.C. (1987) Metaanalyses of randomized controlled trials. *New England Journal of Medicine*, 316:450-455.

	Human Lung Cancer	Mouse Tumor Initiation	Viral Trans- formation	Muta- genesis -MA	Muta- genesis +MA
Roofing Tar	*	*	*	*	*
Coke Oven emis.	*	*	*	*	*
Gasoline fumes		*	*	*	*
Benzo-a-Pyrene			*	*	
Cigarette smoke	*	*	*	*	*
Diesel A		*	*	*	*
Diesel B		*	*	*	*
Diesel C			*	*	*
Diesel D		*	*	*	*
Diesel E	*				

Table Classification of available studies in DuMouchel and Harris (1983). The symbol "*" means that the slope of a dose-response curve, with standard error, is available.

Figure Illustration of the Bayesian model for the prior expectation of the log slopes (dotted lines) and the observed log slopes with 68% error bars, using data corresponding to just three rows and three columns from Table 1. There is clear lack of fit to the lines, indicating the need for an extra component of variance.

ASSESSING, MANAGING AND COMMUNICATING ABOUT THE POSSIBLE RISKS OF EXPOSURE TO 60 Hz FIELDS

M. Granger Morgan
Carnegie Mellon University
Pittsburgh, Pennsylvania

RISK ASSESSMENT

Assessing the possible risks of 60 Hz fields poses some challenging problems to conventional risk assessment. The most serious problem is the fact that, if there is a risk, it is unclear what constitutes dose. While a few of the epidemiological studies suggest that effects increase monotonically with increasing field strength, many of the cellular and animal studies suggest the existence of complex resonant or threshold processes. When a fundamental relationship like this is uncertain, common practices in risk assessment include parameterizing the analysis in order to determine how the results depend upon plausible alternative assumptions about the form of the effects function, or performance of a bounding analysis that, while it does not establish a best estimate for the incidence of health consequences, does establish an upper bound. Attempts to perform both such analyses have given rise to unsatisfactory results. Parametric analysis leads to a combinatorial explosion, with results which span many decades. Bounding analyses set bounds that are very high. Other than saying the impacts could be quite wide spread (because fields are ubiquitous) these results are of little practical utility.

If epidemiological studies continue to produce results that show a monotonically increasing effect with field strength, it may become appropriate to perform estimates using linear effects functions. Two arguments are relevant. The first is that the integrated population impact of a series of exposure functions that involve resonances and thresholds might be approximated by a monotonically increasing function of field strength. The second is that what we see in the epidemiological studies should be viewed as a lower bound estimate. Because most of these studies use field strength as their dose metric, it is quite possible that if they switched to "the correct" metric, the effect they see would become considerably stronger. This line of argument suggests that an assessment based on a monotonically increasing effect function might be viewed as a lower bound. There is a second reason why such an analysis would produce a lower bound estimate. If a linear effects function is used, the incidence base rate in epidemiological studies arguably does not involve unexposed individuals (unless a threshold is assumed), but rather less exposed individuals.

Beyond the problem of not knowing the appropriate definition of "dose," there is great confusion about the possible effects that might result from exposure. Cancer promotion has been the focus of most public discussion. Developmental abnormalities are also a possibility, as are chronic neurological impacts such as depression. However, there is as yet no unambiguous evidence that any such effects occur.

In summary, while it is clear that quantitative risk analysis could produce estimates that involved rather large numbers (i.e., involvement of field exposure in a significant fraction of certain types of cancers, birth defects, etc. all of which have fairly high base rates) given the very incomplete understanding of the science that is now available, I believe that such quantitative estimates would be of little significant use in the policy process, and have the potential of being easily misinterpreted.

RISK MANAGEMENT

While the science is uncertain, the concern it has generated about possible health risks is both real and legitimate. We have identified three possible approaches (1):

1. Conclude that there is not yet enough evidence to warrant any action. Don't make any changes in the way we do things until new research tells us clearly whether there is a risk and, if so, how big it is.
2. Conclude there is some basis for concern. Adopt a position of "prudent avoidance," which means limiting exposures that can be avoided with small investments of money and effort. Don't do anything drastic or expensive until research provides a clearer picture of whether there is any risk and, if there is, how big it is.
3. Conclude we have a real problem and spend some serious time and money on an aggressive program of limiting field exposures now, while recognizing that we may eventually learn that some or all of this effort and money has been wasted, either because it wasn't needed or we spent it the wrong way because we didn't understand the science well enough to spend it effectively.

Prudent avoidance involves doing things which keep people out of fields at modest cost. Using plausible upper bounds on effects frequencies one can estimate costs per effect avoided and compare with social investment rates to avoid known risks. Such arguments make it difficult to justify investments of more than a few thousand dollars per person-exposure avoided (2).

In addition to prudent avoidance, transmission line field strength standards can be justified on two other grounds: equity and safety. Equity arguments say "we don't know if exposure to transmission line fields is more dangerous than exposure to other fields, but we know it is different than the exposures we all normally receive. Thus, it is inequitable. It could be made equitable if exposures could be made "similar" to the exposures people regularly receive from other sources" (2,3). Safety based standards are predicated on being able to say that certain fields will not result in adverse health effects, something which for the range of field exposures of practical interest, current scientific understanding will not support.

Several states have adopted field strength standards for electric fields. Florida has now also adopted such a standard for magnetic fields. While most states have operated with the fiction that they were setting "safe field" standards, several have clearly also had equity considerations in mind. A few have begun to talk explicitly in terms of prudent avoidance. It is clear that such thinking played a significant role in the recently promulgated Florida standards.

RISK PERCEPTION AND COMMUNICATION

Our research on risk communication began with two kinds of scoping studies. Using techniques developed in a large NSF project in risk communication that is now under way in our group, we conducted extended face-to-face interviews with six naive respondents in Pittsburgh, and a group of respondents who had been involved in transmission line siting controversies in upstate New York (8 respondents) and in central Florida (3 respondents). In addition, we conducted three day-long focus groups, two in Pittsburgh with 18 naive respondents and one in upstate New York with 7 respondents who had been involved in a siting controversy. Using a set of protocols developed entirely independently, Paul Slovic and our other collaborators at Decision Research ran two focus groups with 12 naive respondents in Eugene, Oregon.

We undertook these scoping studies with the objective of identifying questions and issues that warranted more detailed examination. Results from these exploratory studies can only be treated as suggestive. They should not be taken as conclusions. However, the fact that the Carnegie Mellon and Decision Research studies were done independently, and led to the same basic conclusions, suggests that what we observed was not random noise.

Our respondents displayed very limited understanding of the nature of 60 Hz electric and magnetic fields. There was confusion about the role of such variables as voltage, current and power, as well as the operation of such process as electrostatic shielding. Understanding of sources and range dependency was also limited. We observed confusion about the differences between 60 Hz fields and ionizing and non-ionizing radiation.

Most of our respondents found the notion that there might be health effects from exposure to fields very plausible. Lists of possible health effects that respondents produced were long and ranged well beyond cancer. Children were perceived as particularly sensitive. Synergisms, as with agents such as defoliant sprays, were suggested as a possibility by many. Respondents displayed great difficulty in producing useful analogies or models. The notion of effects based on the time integral of field strength was seen as compelling, even in the face of countervailing evidence.

Most of our respondents displayed very limited understanding of the process by which science is done. In particular, issues of experimental design, control and statistical significance were poorly understood, hard to grasp and not seen as compelling. We saw substantial confusion about the meaning of correlation and the role of uncertainty. For many of our respondents, results from animal studies were significantly more persuasive than results from cellular studies. Anecdotal evidence was widely seen as very compelling. There was a general perception that getting answers to the question "is there a risk, and if so, how big is it?" should be easy and straightforward.

Many of our respondents, especially those who had been involved in siting controversies, suggested that researchers working in this field were motivated by "interests" rather than by a search for the truth. The source of an investigator's research funds was often seen as significant. A typical reaction to a responsible methodological critique of an epidemiological study was "for God's sake why would you argue against the validity of the study unless you had some sort of position to maintain?" Several respondents in both the Carnegie Mellon and the Decision Research Groups made comparisons to "tobacco company scientists." It is unclear how easily this framing is adopted. Once adopted it is clear that it "explains" a great deal.

Our basic impression from these preliminary studies is that it is not very hard to convince many people that health risks from 60 Hz fields associated with transmission lines are real, are serious, and that regulatory action is warranted. Perceptions of other sources of fields (some involving comparable magnitudes) were less clear, but the level of concern was clearly lower. This is consistent with our earlier findings (4).

Exposure control arguments based on cost effectiveness were found to be hard to grasp and easily misinterpreted. Equity issues were viewed by many as very important. Framings in terms of relative risk were not often undertaken, or seen as especially important. The fact that a lot of attention is being given to this topic was taken as evidence that there is a risk.

Most respondents who had been involved in siting controversies had very strong views about utilities. They displayed little sense of a "public service" role. Utilities were characterized as greedy, arrogant, and uncaring, and these characterizations evoked responses such as punish, force, and compel. Especially among New York respondents, need for new transmission facilities was doubted. Local generation near load centers was viewed by several of our New York respondents as an attractive and viable solution. The other actions of utilities appeared to have had significant impacts on several of our New York respondents' views about the field issue.

While the respondents who have been involved in transmission line siting controversies knew a lot about the issues of 60 Hz fields, our naive respondents had little or no familiarity with the topic. Some of the studies of perception and understanding we plan

can operate with naive respondents. However, for many studies we need to begin by

informing people about fields. In order to do this we have developed a brochure for lay audiences (1). The development of this brochure was substantially informed by the scoping studies outlined above. The choice of some of the topics addressed, and some of the framings of those topics, were affected by these studies. Twelve experts at Carnegie Mellon and Decision Research were involved in selecting and reviewing the technical content and preparing and evaluating the style and presentation. In addition, one lay focus group was devoted entirely to a critique of a draft version of the brochure, and two other lay focus groups were asked to evaluate selected portions of the brochure. In addition, technical content reviews were obtained from fourteen outside experts in a wide variety of fields who are involved in the issues from many different perspectives. Copies of this brochure will be available to participants at the meeting.

Our explanatory studies were completed last fall. Since then we have been engaged in developing and piloting a number of survey instruments and other experiments that are designed to produce more definitive evaluations of most of the same issues. Studies that are furthest along include:

- A questionnaire that explores basic knowledge (and degree of confidence) about the physics of electric and magnetic fields;
- A study that specific understanding about the range dependence of field strength;
- A questionnaire that explores knowledge about and attitudes toward possible field related health effects;
- A video based study that will explore some of the determinants in lay responses to expert testimony. Factors to be examined will include prior knowledge, speaker affiliation, and message content.

A variety of other studies are in earlier stages of design. Detailed descriptions of some of the experiments now in progress, along with a few preliminary findings will be presented in the talk.

At the completion of this work we hope to have the tools to answer the following five questions: (1) What do different publics "know" about power-frequency fields? (2) How do they frame and understand the issues? (3) How does information affect knowledge and framing? (4) How effectively can issues involving the process of science be communicated? (5) What risk management strategies do publics consider appropriate and why?

ACKNOWLEDGEMENT

My principal research collaborators in this work include Connie Cortés, Baruch Fischhoff, Keith Florig, Gordon Hester, Jim Hoburg, Lester Lave, Indira Nair, and Emilie Roth all of Carnegie Mellon and Paul Slovic and Don MacGregor of Decision Research. Support for our work on 60 Hz risk perception and communication is provided under EPRI contract RP 2955-3. General research on risk communication at Carnegie Mellon University is being supported under NSF grant SES-8715564.

REFERENCES

1. M. Granger Morgan, "Electric and Magnetic Fields from 60 Hertz Electric Power: What do we know about possible health risks?" Department of Engineering and Public Policy, Carnegie Mellon University, 1989 March.

can operate with naive respondents. However, for many studies we need to begin by informing people about fields. In order to do this we have developed a brochure for lay audiences (1). The development of this brochure was substantially informed by the scoping studies outlined above. The choice of some of the topics addressed, and some of the framings of those topics, were affected by these studies. Twelve experts at Carnegie Mellon and Decision Research were involved in selecting and reviewing the technical content and preparing and evaluating the style and presentation. In addition, one lay focus group was devoted entirely to a critique of a draft version of the brochure, and two other lay focus groups were asked to evaluate selected portions of the brochure. In addition, technical content reviews were obtained from fourteen outside experts in a wide variety of fields who are involved in the issues from many different perspectives. Copies of this brochure will be available to participants at the meeting.

Our explanatory studies were completed last fall. Since then we have been engaged in developing and piloting a number of survey instruments and other experiments that are designed to produce more definitive evaluations of most of the same issues. Studies that are furthest along include:

- A questionnaire that explores basic knowledge (and degree of confidence) about the physics of electric and magnetic fields;
- A study that specific understanding about the range dependence of field strength;
- A questionnaire that explores knowledge about and attitudes toward possible field related health effects;
- A video based study that will explore some of the determinants in lay responses to expert testimony. Factors to be examined will include prior knowledge, speaker affiliation, and message content.

A variety of other studies are in earlier stages of design. Detailed descriptions of some of the experiments now in progress, along with a few preliminary findings will be presented in the talk.

At the completion of this work we hope to have the tools to answer the following five questions: (1) What do different publics "know" about power-frequency fields? (2) How do they frame and understand the issues? (3) How does information affect knowledge and framing? (4) How effectively can issues involving the process of science be communicated? (5) What risk management strategies do publics consider appropriate and why?

ACKNOWLEDGEMENT

My principal research collaborators in this work include Connie Cortés, Baruch Fischhoff, Keith Florig, Gordon Hester, Jim Hoburg, Lester Lave, Indira Nair, and Emilie Roth all of Carnegie Mellon and Paul Slovic and Don MacGregor of Decision Research. Support for our work on 60 Hz risk perception and communication is provided under EPRI contract RP 2955-3. General research on risk communication at Carnegie Mellon University is being supported under NSF grant SES-8715564.

REFERENCES

1. M. Granger Morgan, "Electric and Magnetic Fields from 60 Hertz Electric Power: What do we know about possible health risks?" Department of Engineering and Public Policy, Carnegie Mellon University, 1989 March.

2. M. Granger Morgan, H. Keith Florig, Indira Nair, and Gordon L. Hester, "Controlling Exposure to Transmission Line Electromagnetic Fields: A regulatory approach that is compatible with the available science", Public Utilities Fortnightly, pp. 49-58, 1988 March.
3. H. Keith Florig, "Population Exposures to Power-Frequency Fields: Concepts, components and control," Ph.D. Thesis, Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA, 1986 November.
4. M. Granger Morgan, Paul Slovic, Indira Nair, Dan Geisler, Donald MacGregor, Baruch Fischhoff, David Lincoln, and H. Keith Florig, "Powerline Frequency Electric and Magnetic Fields: A pilot study of risk perception," Risk Analysis, 5, pp. 139-150, 1985 June.

ELECTRIC FIELDS: DISCUSSION AND SUMMARY **(Does the shape of the shovel really matter?)**

Michael E. Ginevan
Risk Focus Division, Versar, Inc.
Springfield, Virginia

INTRODUCTION

At the start of the conference I offered a set of questions that was meant to summarize what we hoped to accomplish this week:

1. What is the state of our knowledge of potential health effects of electric fields, and what sorts of quantitative approaches have been used to date?
2. How might new statistical theory be used to improve the design and conduct of studies to address problems where, as is the case for electric fields, the initial hypotheses are diffuse?
3. Can statistical methodologies such as quantitative risk assessment, meta-analysis, and decision theory be used to obtain improved risk estimates from existing data?
4. What is the appropriate interaction between biological and statistical experts in such endeavors?
5. When do we have enough evidence to begin intensive, directed study of a problem?
6. When do we have enough evidence to stop studying a problem?
7. How can we use data collected to date to best guide the ongoing research program?
8. What constitutes clear evidence of a hazard?

The last question is how successful have we been in addressing these issues? Overall, I think we have addressed many if not most of them. I know I learned a lot about the biology, and how statistics might fit into future studies.

In my discussion, I will stress three points. First, I would like to reprise what I think I learned about the biology. Second, I will play statistical consultant, and go through some of the statistical approaches that I think might help those studying electric fields. Finally, I will close with a couple of thoughts on policy issues that will get us back to Granger Morgan's presentation.

BIOLOGICAL ISSUES

In the biological arena, it seems that three main areas can be identified, epidemiology (cancer), cellular studies, and neuro-behavioral studies. These areas have some logical points of interaction. Cellular studies can contribute mechanistic understandings that will enhance the conduct of both epidemiology and neuro-behavioral studies. I think it is questionable, in a scientific sense, what epidemiology studies have to say to neuro-behavioral studies, so if I were asked to assess research needs, I would say there are two sets of separable issues. I think in the problem area, I've got three conclusions:

1. Epidemiology studies are pretty much separate from neuro-behavioral studies, because epidemiology studies focus almost entirely on cancer endpoints, which presumably have very different mechanistic bases from neuro-behavioral endpoints.
2. From 1, cellular studies conducted in support of epidemiology are likely different in focus from cellular studies conducted in support of neuro-behavioral problems.
3. If you are looking for mechanisms, doing phenomenological studies, which is really what epidemiology studies are, will not identify them. Both laboratory and field studies are necessary and are best conducted by investigators working in concert.

None of these insights are very profound, but may be somewhat useful in organizing our thinking regarding the conduct and interpretation of electric field studies. For example, you cannot really talk about positive epidemiology studies as lending support to neuro-behavioral results because of the probable difference in mechanistic basis.

STATISTICAL ISSUES

One central question of the conference was how statisticians might improve the conduct of electric fields health effect studies. Rather than try to summarize the preceding presentations on this topic, I'd like to try and illustrate how such aid might be delivered by a consulting statistician. The first question is, what is the appropriate role for a consulting statistician? The first step in defining a role is realizing what a statistician cannot do for you. This was put forth earlier as Chiacchierini's inequality:

Statistician = Magician.

That is, if you want a statistician to quickly pull an answer out of a hat for you, you are likely to be disappointed.

Turning to what a statistician can do, I have always liked Donald Marquardt's (1987) role definition: A statistician is a "purveyor of the scientific method." Starting with this viewpoint, I would like to walk through an example of how a statistician might contribute to a research program designed to study cancer risks related to electric field exposure.

The first observation I will offer is that such a study is really concerned with magnetic field exposure. The reason for this is that we know that positive results have been shown in terms of people living close to powerlines having elevated cancer risk. Now, we have seen that, while magnetic field exposure is fairly readily predictable from powerline configurations, electric field exposure is not. From my training as a statistician, I know that if I am looking for an association between some environmental exposure and some disease risk, I must have my study population well classified with respect to dose. If I do not, even real effects will be obscured. Since effects have been found (assuming they are not artifact of some kind) the study populations must be fairly well classified with respect to dose. Thus, it follows that the culprit must be magnetic as opposed to electric fields.

So, what do we know about magnetic fields? First, we can say with confidence that they are not genotoxic (this may be one of John Bailer's "anchors"). Thus, they are likely to be tumor promoters (if indeed, they do anything). We obviously do not know how they act, but a statistician can, I think, aid speculation. In Figures 1 and 2, I have drawn two hazard functions (non-statisticians: this is approximately equivalent to the cancer rate over time) that show how electric field exposure might influence the hazard. In the first, when exposure commences, the hazard increases rapidly, plateaus at a new level (which

might be proportional to "dose") and then rapidly returns to its original level when exposure ceases. In the second, the hazard rises throughout the period of exposure and stays at the new higher level when exposure ceases. Both of these models are doubtless wrong, but they illustrate how a statistician can throw out models, which can be criticized and tested by biological investigators, which in turn stimulates improved models. If enough of this sort of interaction can be sustained real progress may ensue.

What else can a statistician do? Well, he can suggest pieces of information that would seem helpful in quantifying the behavior of a given system. It seems, for example that modification of the electric environment of a cell, or more probably a tissue, is what must modify the cancer hazard. Thus, I would like to have information on the electrical environment of normal tissues and how this environment is modified when a malignancy occurs. I would also like some information on how the electrical environment if a tissue is modified by electric fields, so that this change can be compared to both the normal and malignant cases. Having obtained this information, a statistician can help develop models based on it (e.g., how a modified electrical environment might foster the development of cancer), which in turn will suggest new studies.

A statistician can also help design studies. Say we are interested in a field epidemiology study. A statistician can help identify a target population (which must be sufficiently exposed and numerous). He can also assist in addressing such questions on whether stratification or matching is best in a particular study, and if an a priori guess as to effect size is available can determine the study's power (its ability to detect real effects of specified magnitude), and attendant questions of required sample size. Last (and by far least; I would guess that most statisticians spend less than 5% of their time analyzing data) the statistician can analyze the data once it is collected.

One last point on statistics is that, in the sorts of prospecting studies carried out in a diffusely defined research area like potential health effects of electric fields, a Bayesian perspective is very helpful. DuMouchel showed us an excellent example of the problems that can arise in a frequentist approach to multiple comparisons. Because many comparisons are involved, we lengthen the confidence intervals on each effect measure, which results in raising the upper bound on the largest effect, which the theory order statistics tells us is larger than its true expectation to start with! Still, in areas such as environmental studies, where multiple comparisons are the rule, not the exception, I feel uncomfortable with an approach that does not penalize me in some way for all the comparisons I have made. Perhaps the answer can be found in something like Thrall's suggestion of a multivariate prior distribution. In any case, I think this highlights an important area for further statistical work.

POLICY ISSUES

My final comments are, as promised, in the policy arena. Figure 3 displays something I'll call "Policy Space." Which has 3 dimensions: quality, relevance, and perception. As Morgan has shown here, a study with X,Y,Z scores of 0,0,1 (all axes are unit length) is likely to be very influential in the arena of public opinion, despite the fact that it is irrelevant and of miserable scientific quality. There is a real challenge here in the arena of public education. We have to make the public aware of what a good study is, and really, what scientists actually do, so that when confronted with snake oil they identify it as such.

Finally, I'd like to consider policy decisions. As Bailar pointed out, we need stopping rules. To this I'd like to add pruning rules and starting rules. The difference between stopping rules is that a stopping rule says "this really is not a worthwhile line of research and should be stopped," while a pruning rule says "this is good research, but it is not relevant to our program on electric fields health effects any longer, and cannot be funded out of this pot any longer." My observations suggest that the latter may be at least as important as the former. Many research programs whose central focus is the delineation of human risk seem to me to be unduly burdened by excellent research that is almost

entirely irrelevant to the presumed mission of the program (for obvious reasons, I will refrain from offering examples). The notion of starting rules is something I threw out earlier in the conference. On reflection, I realize that it is hard to come up with an example of a study that had political support, but that was not started simply because it was a bad idea. Still, firm rules for not starting poorly conceived or simply impossible studies could be worthwhile.

CONCLUSIONS

In closing, I'd like to address the parenthetical element of my title. The shape of the shovel (our methodologies and approaches) matters very much indeed. In addressing difficult issues like electric fields, a lot of interdisciplinary collaboration is a must. The approaches of biological investigators must be shaped in part by statistical considerations, and the approaches of statisticians must reflect biological reality and unusual features of exploratory research projects such as multiple comparison problems. We must also be aware that when research is conducted in support of decision making on environmental issues, we as scientists have an obligation to really educate the lay public, and should be ready, as "experts" to enter the public policy arena ourselves. I know this conference has helped me to focus on how biologists and statisticians might better collaborate, and has caused me to examine many issues I might have otherwise not. In short, it has been a real learning experience for me, and I hope that it has provided similar rewards for everyone.

REFERENCES

Marquardt, D.W. 1987. The Importance of Statisticians. Journal of the American Statistical Association 82: 1-7.

Figure 1. A hypothetical hazard function for electric field exposure, here the hazard fades after exposure stops.

Figure 2. A hypothetical hazard function for electric field exposure, here the elevated hazard persists after exposure stops.

Figure 3. A three dimensional view of "Policy Space."

**APPENDIX I
CONFERENCE PROGRAM**

SUNDAY, JULY 9

TUTORIALS

Chair: *Michael E. Ginevan*, RiskFocus, Inc.

- 4:00 p.m. **Introduction to Electric and Magnetic Fields**
Thomas S. Tenforde, Pacific Northwest Laboratory
- 5:00 p.m. **Introduction to Statistics**
Richard P. Chiacchierini, Food and Drug Administration

MONDAY, JULY 10

- 9:00 a.m. **Why Are We Here?**
Michael E. Ginevan, RiskFocus, Inc., Chair, 8th Conference on Radiation and Health
- 10:00 a.m. **EPIDEMIOLOGICAL REVIEW I**
Chair: *John C. Bailer*, McGill University and U.S. Public Health Service
- Studies of Cancer and Residential Exposure**
Charles Poole, Epidemiology Resources, Inc.
- Case Control Studies: Frequentist vs. Bayesian Points-of-View**
Marvin Zelen, Harvard School of Public Health
- Discussant: *Emanuel Landau*, American Public Health Association
- Floor Discussion
- 7:30 p.m. **EPIDEMIOLOGICAL REVIEW II**
Chair: *Peter G. Groer*, Oak Ridge Associated Universities
- Occupational Studies**
Gilles Theriault, McGill University
- Models of Exposure in Retrospective Cohort Studies**
Jack Sahl, Southern California Edison
- Discussant: *Robert Goldsmith*, U.S. Department of Energy
- Floor Discussion

TUESDAY, JULY 11

- 9:00 a.m. **EXPERIMENTAL STUDIES I**
Chair: *Anthony D. Thrall*, Electric Power Research Institute
- A Summary of Health Research Related to 60 Hz Electromagnetic Fields**
Lee R. Rosen, W/L Associates
- Human Laboratory Studies on the Effects of 60 Hz Fields**

Discussant: *Michael E. Ginevan*, RiskFocus, Inc.

Floor Discussion

7:30 p.m.

EXPERIMENTAL STUDIES II

Chair: *Mays L. Swicord*, Food and Drug Administration

Exposure Assessment for Power Frequency Electric and Magnetic Fields
T. Dan Bracken, T. Dan Bracken, Inc.

Statistical Issues in Exposure Assessment
Leeka I. Kheifets, Electric Power Research Institute

Discussant: *Richard P. Chiacchierini*, Food and Drug Administration

Floor Discussion

WEDNESDAY, JULY 12

9:00 a.m.

SPECIAL PROBLEMS I

Chair: *Lee R. Rosen*, W/L Associates

What is Dose?
Mays Swicord, Food and Drug Administration

Effect of Dosimetry Errors
Clay Easterly, Oak Ridge National Laboratories

Discussant: *Peter G. Groer*, Oak Ridge Associated Universities

Floor Discussion

7:30 p.m.

SPECIAL PROBLEMS II

Chair: *Emanuel Landau*, American Public Health Association

Multi Comparisons
A. A. Afifi, UCLA School of Public Health

Statistical Issues in Exploratory vs. Confirmatory Research
Anthony D. Thrall, Electric Power Research Institute

Discussant: *Richard L. Anderson*, Statistical Consultants, Inc.

Floor Discussion

THURSDAY, JULY 13

9:00 a.m.

PANEL ON RESEARCH STRATEGIES

Chair: *Peter G. Groer*, Oak Ridge Associated Universities

Panelists:

Kay T. Kimball, Baylor College of Medicine
Meta-Analysis

John C. Bailar, McGill University and U.S. Public Health Service
Stopping Rules

William H. DuMouchel, BBN Software Products
Bayesian Methods for Combining Dose-Response Studies

Question and Answer Period

7:30 p.m.

RESEARCH POLICY AND RISK ASSESSMENT
Chair: *Michael E. Ginevan*, RiskFocus, Inc.

**Assessing, Managing, and Communicating About the Possible Risks of
Exposure to 60 Hz Fields**
M. Granger Morgan, Carnegie Mellon University

Floor Discussion

Conference Summary
Michael E. Ginevan, RiskFocus, Inc.

**APPENDIX II
LIST OF PARTICIPANTS**

A. A. Afifi

UCLA School of Public Health
10917 Barman Avenue
Culver City, CA 90230
(213) 825-6381

Richard L. Anderson

Statistical Consultants, Inc.
Quality Place, 300 East Main St., #400
Lexington, KY 40507-1539
(606) 252-3890

Barbara A. Bailar

American Statistical Association
1429 Duke Street
Alexandria, VA 22314
(703) 684-1221

John C. Bailar III

468 N Street, S.W.
Washington, DC 20024
(202) 472-5584

Robert S. Banks

Robert S. Banks Associates, Inc.
825 Washington Avenue, S.E., #104
Minneapolis, MN 55414
(612) 623-4646

Michael Barge

Locus, Inc.
P.O. Box 740
State College, PA 16804
(814) 466-6275

T. Dan Bracken

T. Dan Bracken, Inc.
5515 S.E. Milwaukie Avenue
Portland, OR 97202
(503) 233-2181

Bobby G. Burgess

Houston Lighting and Power Company
P.O. Box 1700
Houston, TX 77251
(713) 623-3702

John S. Campbell

Salt River Project
P.O. Box 52025
Environmental Services Department
Phoenix, AZ 85072-2025
(602) 236-2008

Richard P. Chiacchierini

U.S. Food & Drug Administration
5600 Fishers Lane (HFZ-116)
Rockville, MD 20857
(301) 443-7120

Mary R. Cook

Midwest Research Institute
425 Volker Avenue
Kansas City, MO 64110
(816) 753-7600, ext. 162

William E. Cory

Cory Consultants, International, Inc.
4435 High Sierra
San Antonio, TX 78228
(512) 736-0714

Fred M. Dietrich

Electric Research and Management, Inc.
2140 William Pitt Way
Pittsburgh, PA 15238
(412) 826-3222

Necip Doganaksoy

Union College
1057 University Place
Schenectady, NY 12308
(518) 377-5448

Tom Downs

University of Texas
School of Public Health
P.O. Box 20186
Houston, TX 77025
(713) 792-4421

Christiana M. Drake

University of Washington
Department of Biostatistics
2423 Queen Anne Avenue North
Seattle, WA 98109
(206) 545-2737 or 223-5421

Thanh Doan

System Development Division
Power Grid Development Department
State Electric Commission of Victoria
15 William Street
Melbourne 3000, Australia

William H. Du Mouchel
BBN Software Products Corporation
10 Fawcett Street
Cambridge, MA 02138
(617) 873-8119

Clay Easterly
Oak Ridge National Laboratory
P.O. Box 2008, M.S. 6101
Oak Ridge, TN 37831-6101
(615) 574-6254

Andrew E. Gelman
Harvard University
Department of Statistics
Science Center, 6th Floor
1 Oxford Street
Cambridge, MA 02138
(617) 495-5325

Mary Gerkovich
Midwest Research Institute
425 Volker Boulevard
Kansas City, KS 64110
(816) 753-7600, Ext. 164

Michael E. Ginevan
RiskFocus, Incorporated
6850 Versar Center
Springfield, VA 22151
(703) 750-3000, ext. 521

Carl V. Gogolak
U.S. Department of Energy
Environmental Measurements Laboratory
376 Hudson Street
New York, NY 10014
(212) 620-3635

Robert Goldsmith
U.S. Department of Energy
Office of Health and Environmental
Research, ER73-GTN
Washington, DC 20545
(202) 353-4739

Peter G. Groer
Oak Ridge Associated Universities
P.O. Box 117
Oak Ridge, TN 37831-0117
(615) 576-3532

Ronald B. Harrist
University of Texas
School of Public Health
P.O. Box 20186
Houston, TX 77225
(713) 792-4480

Donald E. Herbert
University of South Alabama
College of Medicine
Department of Radiology, 269 CSAB
Mobile, AL 36688
(205) 460-7064

Doreen Hill
U.S. EPA
1366 Moyer Road
Annapolis, MD 21403
(202) 475-9626

Marilyn J. Humm
American Statistical Association
1429 Duke Street
Alexandria, VA 22314
(703) 684-1221

Harriet Keener
Martin Marietta Energy Systems
Building 9711-1, MS-8105
P.O. Box 2009
Oak Ridge, TN 37831
(615) 574-3538

Michael Kelsh
UCLA
18548 Fidalgo
Rowland Heights, CA 91748
(818) 444-8098

Leeka I. Kheifets
Electric Power Research Institute
3412 Hillview Avenue
PO Box 10412
Palo Alto, CA 94303
(415) 855-8976

Kay T. Kimball
Baylor College of Medicine
Design and Analysis Core/Med.
6550 Fannin Street, SM-1201
Houston, TX 77030
(713) 798-4213

Emanuel Landau
American Public Health Association
4601 N. Park Avenue, #208
Chevy Chase, MD 20815
(301) 789-5685

Patrick M. D. Levallois
DSC-Center Hospitalier de
l'Université Laval
2050 Bld. St. Cyrille
Ste-Foy, PQ, Canada G1X 3W6
(418) 687-1090

M. Granger Morgan
Carnegie Mellon University
Department of Engineering and
Public Policy
Pittsburgh, PA 15213
(412) 268-2672

Jacques Ortolan
TDF of France
2127 Barbee
Mt. Rouge 92120, France
49-65-1202

Bertil Persson
Lund University
Radiation Physics Department
Lasarettet
S-22185 Lund, Sweden
(46) 46-173111

Charles Poole
Epidemiology Resources, Inc.
826 Boylston Street, #200
Chestnut Hill, MA 02167
(617) 734-9100

James F. Prewitt
SRI International
1900 Randolph Road, S.E.
Albuquerque, NM 87106-4245
(505) 842-0961

Lee Rosen
National Institutes of Health
5333 Westbard Avenue, Rm. 2A10
Bethesda, MD 20892
(301) 496-7223

Jack D. Sahl
Southern California Edison Company
2244 Walnut Grove
Rosemead, CA 91770
(818) 302-9696

Nuala A. Sheehan
University of Washington
5012 20th Avenue N.E.
Seattle, WA 98105
(206) 543-8265

Mays Swicord
Food and Drug Administration
Div. of Life Sciences, HRZ-113
5600 Fishers Lane
Rockville, MD 20857
(301) 443-7153

Thomas S. Tenforde
Pacific Northwest Laboratory
K-4-14, P.O. Box 999
Richland, WA 99352
(509) 375-3738

Gillis Theriault
School of Occupational Health
McGill University
1130 Pine Avenue West
Montreal, PQ, Canada H3A 1A3,
(514) 398-4227

Anthony D. Thrall
Electric Power Research Institute
3412 Hillview Avenue
PO Box 10412
Palo Alto, CA 94303
(415) 855-2627

Howard Wachtel
University of Colorado
and Electric Power Research Institute
Box 425
Boulder, CO 80309-0425
(303) 492-7713

Deborah Weil
ERI, Incorporated
575 Eighth Avenue
New York, NY 10018-3011
(212) 465-1060

J. Ken Wiley
Florida Electric Power Coordinating Group
402 Reo Street, #214
Tampa, FL 33609
(813) 289-5644

Marvin Zelen
Dept. of Biostatistics
Harvard School of Public Health
677 Huntington Avenue
Boston, MA 02115
(617) 732-1056

ASA Coordinator
Lee L. Decker
American Statistical Association
1429 Duke Street
Alexandria, VA 22314
(703) 684-1221

**DATE
FILMED**

2 / 18 / 93