CONF-9107220--

DE92 014343

REPORT ON THE SECOND INTERNATIONAL WORKSHOP ON RESIDENTIAL RADON

U.S. Department of Energy
Office of Energy Research
Office of Health and Environmental Research
Washington, D.C. 20585

Commission of European Communities
Radiation Protection Programme
Brussels, Belgium

July 1991
Alexandria, VA

Radon

MASTER

Workshop Proceedings

FOREWORD

Since the discovery in 1984 of a house in Pennsylvania with several thousand picocuries per liter of radon indoors, other areas of the country have been found with significantly elevated radon levels. In this same period of time the issue of public health risk from indoor radon has grown in both importance and controversy. Significant scientific questions yet remain about extrapolating the lung cancer risk from radon exposure in uranium miners, the source for radon risk estimates, to that of the general public. Residential radon epidemiology is viewed as the most direct approach to resolving these uncertainties in extrapolation of risk, and many such studies are planned or underway both in the United States and abroad.

In early 1988, it was still unclear as to how many of these studies were being performed and more importantly how comparable the study designs might be. The Department of Energy (DOE), being the lead agency for radon related basic research, was interested in the number, location, and design of these studies. At that time, Dr. John S. Neuberger, of the University of Kansas began a small task for DOE to identify residential radon epidemiology projects world-wide and to gather basic information about each. DOE utilized Dr. Neuberger's findings to convene a workshop of all the active residential radon epidemiologists to evaluate the status of these studies.

The combined efforts of DOE, the Commission of European Communities (CEC), Dr. Jonathan Samet, University of New Mexico, Dr. Jan Stolwijk, Yale University School of Medicine, and Dr. Neuberger culminated in a meeting in Alexandria, Virginia, in July 1989, where 18 participating scientists, and many more observers, representing ten countries convened. The agenda was designed to establish a consensus on such things as the need, if any, for further epidemiology studies, the interest in a major data pooling effort, and the possibility of designing comparable protocols.

The present report summarizes the results of a second meeting, convened 2 years after the original. It was designed to reestablish contact between the investigators, encourage the gathering of common data, address outstanding problems, resolve exposure issues and begin the strategy for pooling eventual study results. This second successful meeting accomplished these goals. Both the DOE and the CEC are now formulating plans to implement the consensus recommendations of the scientists involved.

Further information about the Radon Research Programs can be obtained from:

DOE Program

Dr. Susan L. Rose
Radon Program Manager
Office of Health and Environmental Research, ER-72
Department of Energy, GTN
Washington, DC 20585

CEC Program

Dr. Jaak Sinnaeve
Commission of the European Communities
Nuclear Safety Research/Radiation Protection Programme
Rue de la Loi 200
B-1049 Brussels, Belgium

TABLE OF CONTENTS

*Reprinted with Permission

EXECUTIVE SUMMARY

As a follow-on to the First International Workshop on Residential Radon Epidemiology held in Alexandria VA, on July 24-26, 1989, a Second Workshop was convened, in Alexandria, VA, July 22-23 1991, also under the auspices of the U.S. Department of Energy and the Commission of European Communities. The Workshop, co-chaired by Jonathan Samet and Jan Stolwijk, was attended by 20 active participants from seven countries representing epidemiologic studies recently completed, currently in progress, or in the last stages of preparation. The studies reported on are being conducted in the United States, Canada, Sweden, the United Kingdom, France, Belgium, Germany and the Peoples' Republic of China.

The invited presentations that initiated the Workshop focused on a number of methodological problems that have surfaced in the last few years. Among these were:

1. the difficulties in predicting indoor radon concentrations, based on geologic information,(discussed by Alan Tanner, formerly of the U.S.Geologic Survey);

2. the relationships between indoor radon concentrations and building characteristics (discussed by Richard Sextro, Lawrence Berkeley Laboratory, U.S.A.);

3. the approaches to analysis of case-control studies in radon epidemiology (discussed by Sarah Darby, Imperial Cancer Research Fund, UK);

4. statistical approaches to error in measurements and missing data (discussed by Donna Spiegelman, Tufts University, U.S.A.);

5. preliminary results of a data pooling effort dealing with several different studies of residential radon epidemiology and the lessons to be drawn from this effort (discussed by Jay Lubin, U.S. National Cancer Institute).

Project Status Reports

The meeting participants presented brief reports on the status of the different studies, including projections of the rate of progress, and of dates at which study results might become available for pooling of data.

The participants considered in some detail the optimal ways in which pooling from the currently ongoing studies might best be accomplished. For these discussions the workshop was divided into two groups; one of these groups considered the statistical approaches to pooled data analysis and the optimal way of accomplishing the analysis of the pooled data. The second group considered the data to be contributed to the pooled analysis, the date(s) when such data might be available and how decisions should be made on including or excluding data from the pooled analysis.

Consensus

A consensus was reached on critical issues to be considered in more detail, preferably on a continuing basis, so that the decisions required during the conduct of the studies, and decisions about data management in each of the study centers could be coordinated. If standardization can be achieved during the data collection and the data management phases of the studies then the preparation for pooled analysis could be greatly simplified.

In order to prepare for and facilitate the pooling effort the participants recommended that two focal points for communication and coordination be set up, one in the U.S. and one in the European Region, to be located at a study center. These focal points should be staffed with at least one half-time-equivalent in the near future, until the end of the pooling effort. This person would work to assure coordination and uniformity among the centers.

There was universal agreement that the pooling effort was a worthwhile activity, and that all the investigative teams for the studies represented at the Workshop would participate.

PRESENTATIONS: COMMON PROBLEMS IN RESIDENTIAL RADON EPIDEMIOLOGY

All studies of the epidemiology of lung cancer and residential radon exposure have certain problems in common. A major problem likely to affect all studies is missing data in the characterization of residential radon concentrations. Cases and controls tend to have lived in a number of residences, and although the concentration in their current residence can be obtained with reasonable accuracy, this is not always true for past homes.

If the past homes can be ascertained and identified, and the current occupant and owner consent to placement of a measuring device, the resulting measurement reflects the current conditions in that residence, and it may be necessary to adjust that measurement on the basis of known structural or other changes made in the residence since the case or the control occupied that house.

Some of the past residences cannot now be studied because they have been destroyed, or because the current owner or occupant cannot be persuaded to permit radon concentration measurements to be made. As a result all the studies will be affected by a substantial problem of missing radon exposure data. If estimations are to be made it is important to understand the factors determining the residential radon concentration so that the best possible estimate of exposure can be made, based on whatever data can be collected. The two major groups of factors determining indoor concentrations of radon are those related to the source strength of the underlying soil, and to the characteristics of the residence affecting the exchange of outdoor with indoor air.

"Possibilities and Caveats for using Geological Data to Fill in Missing Exposure Data" discussed by **Dr. Alan Tanner** (formerly of the U.S.G.S.).

Numerous investigations have shown that there is a correlation between radon concentrations in houses and the geological characteristics of the soil and the rock on which the houses are sited. The correlations are good where the indoor radon concentrations are greater than action levels. The basic processes that permit radon entry into a house are known. The efficiency of radon entry into a house is primarily dependent on the radium concentration in the soil and rock within a meter or two of the structure, on the degree of water saturation of the soil, on the soil's permeability to flow of radon-bearing gas, and on the structure's receptiveness to radon entry. Although these primary soil characteristics are in general not directly available in data bases, cross-correlations with mapped or tabulated characteristics can often be expected.

"Predictive" techniques are being developed for estimating the indoor radon potential of the ground at various scales: national, state or provincial, regional, county, and site-specific. So far these approaches have proved fairly successful. These techniques evaluate different combinations of characteristics, depending on the scale of the radon-mapping effort. For

instance, production of a national map relies heavily on aerial surveys by means of detectors of the gamma radiation emitted by one of the radon decay products from the upper 25 cm of the ground, on the physical and radiometric characteristics of physiographic provinces and major rock types, and on the climatic regimes that are typical for major areas; at a county scale (1000 km2), radiometric data are seldom available, but much more specific data on permeability and drainage can be considered. One map of radon potential for U.S. counties showed significantly different, probably log-normal, frequency distributions of indoor radon measurements for the various mapped radon-potential zones. Site-specific ground testing can be done to measure the soil characteristics directly, leading to an estimate of the indoor radon concentration at the time of the test; developments are needed to adjust the results to typical or average conditions, since soil moisture in particular is highly variable. Site-specific testing is not often done because the results if they indicate low radon potential, do not guarantee that mitigation will be unnecessary, and the testing of a single lot may cost nearly as much as mitigation of the house to be built on it. However, if a good estimate of the radon potential of a site - even if the house is no longer there - is of high value (roughly US$ 1000), then site-specific methods are worth considering.

Radon potential assessments at scales other than site-specific can be expected to yield a geometric mean indoor radon potential with a geometric standard deviation. The accuracies of both parameters depend on the quality, density, and appropriateness of the data which are usually surrogates, and on the talents of the persons doing the assessment. Translation of the potential to actual indoor radon values requires that the typical coupling between the ground and structures be known in order to furnish an overall estimate and geometric standard deviation for the indoor radon at a particular location. Although such a degree of uncertainty might be acceptable for a cohort study wherein many samples were drawn from the area for which the radon potential was assessed, the uncertainty might be unacceptably great for a case-control study.

One mistake made in some epidemiological studies has been to use bedrock geological maps, often of an inappropriate scale, to provide a "geological" control. Bedrock maps show what the rock type is at the shallowest depth where it can be found. Above the bedrock may be many meters of rock that have decomposed, altering its characteristics significantly, and several soil layers. The soil may even have been derived from a different rock type and transported to its present location. Because radon's average migration distance is usually limited to one or two meters, the characteristics of bedrock at greater depth are much less important than those of the soil, and sometimes irrelevant.

In conclusion it could be stated that standard geological maps are not usually appropriate for indoor radon potential prediction, and specifically radon potential maps of the wrong map scale will give inaccurate information. Unless cases come from a small neighborhood, only site-specific determinations of radon potential should be valid for case-control studies.

"Structural and climatic factors of importance in determining the residential radon concentration" were discussed by **Dr. Richard Sextro** of the Lawrence Berkeley Laboratory.

There are many structural and climatic factors that enter into this determination. The strength of the radon source at the site is of great importance, but seasonal and meteorologic factors play a role. The tightness of the structure where it touches the soil is important, as well as the tightness of the building skin. Differential air pressures inside the building, and differential air pressures between the inside and outside of the building play an important role. Even with knowledge of a large number of the important determinants of concentration it is not possible to estimate the residential radon concentration with satisfactory accuracy. When estimates are made with the help of multiple regression models it appears that at most 20% of the total variation in radon concentration from one residence to another can be accounted for by the regression model. It was therefore argued that the estimation of residential radon concentration on the basis of building characteristics alone was not likely to produce usable estimates.

The limitations on making estimates of missing data in the residential radon concentration using geology and building characteristics were presented as severe. As a result the epidemiologic studies in this area are likely to have to develop their own approaches to the problem of missing radon data. It was pointed out that in the course of conducting these studies the data gathered on housing characteristics and on radon concentrations in the defined study areas will represent more thorough coverage in terms of radon measurements than was available before, and that extrapolation from that residential radon concentration database may represent the best approach for the missing data problem in each of the studies.

"Analysis of Case-Control Studies" was presented by **Dr. Sarah C. Darby** (Imperial Cancer Research Fund).

Although there is general agreement that the basic approach should consist of stratifying the data, and then performing conditional logistic regression there are many difficulties associated with the execution of a particular analysis.

Having first decided what the appropriate criteria for inclusion or exclusion should be, the next step is to find an appropriate degree of stratification, including effects of interviewer and of large-scale geographic regions, if necessary. The appropriate classification scheme will vary from study to study depending on both the study design and local conditions.

The most important potential confounding variable is clearly cigarette smoking, and great care is needed to make sure that this is appropriately modelled in the analysis. In particular it is inappropriate to assume that the effect of smoking a given number of cigarettes will be the same in different countries. Care should be given to ensure that any confounding effect of other variables is fully taken into account.

Some special problems with studies of indoor radon arise from the difficulties associated with estimating exposure accurately. These include adjustment for seasonal variation if the measurements have not been made over a full year, adjustment for effects such as the installation of double-glazing, weighting the radon concentrations in different years appropriately, and estimating radon concentrations where no measurement can be made. Further methodological work on this latter issue would be desirable.

"Approaches to Exposure measurement error and misclassification" were addressed by **Dr. Donna Spiegelman** (Tufts University).

Two general strategies are most applicable in the setting of residential radon epidemiology, and the uses and limitations of each were discussed:

1) sensitivity analysis such as using theoretical principles such as "non-differential measurement error biases the measure of effect towards the null" to infer the probable effects of measurement error on observed study results and

2) the design of main study / validation study schemes so that effect estimates can be explicitly corrected for bias in the analysis. The latter approach is favored, and several papers were referenced which give guidance on how to determine cost-efficient main study / validation study sizes.

Common models for disease incidence and exposure measurement error and misclassification were reviewed. Preliminary measurement error models for average current radon exposure and for cumulative radon exposure were suggested. Current methods for point and interval estimation of relative risk which correct for bias due to measurement error were presented. Further research is needed to develop methodology to handle measurement error in cumulative exposure variables, and which are appropriate and computationally tractable in a more general class of case-control studies than are currently available.

"Preliminary Results of Data Pooling" presented by **Dr. Jay Lubin** of NCI.

Some very preliminary results of data pooling based on some of the studies represented among the participants were presented. The studies involved were the New Jersey study published by Janet Schoenberg et al., the Swedish study by Goran Pershagen et al. and the study conducted in the Peoples' Republic of China by Blot et al. It was acknowledged that a considerable effort had to be expended in manipulating the data to achieve a reasonably common base of expressing the observations. The preliminary results as presented aroused considerable discussion and it became clear that as an example of data pooling this study identified potential difficulties which might arise in such a pooling effort, especially if the pooling effort included a re-analysis of individual study data. The data manipulation to achieve a homogeneous data set might be more effectively done at the individual study sites where the data were collected.

Working Group Sessions.

Following these presentations the Workshop divided itself into two subgroups: the first subgroup devoted itself to issues having to do with the statistical analysis of pooled data. This subgroup was led by Sarah C. Darby, with Colin Muirhead (NRPB, UK) as the rapporteur. The second subgroup led by Jan Stolwijk concerned itself with issues of data sharing, and the time frame within which data from various studies would be available. There was also discussion of methods to develop reasonably similar data structures between studies so that data management tasks in the pooled analysis would be kept to reasonable levels, hopefully lower than in the preliminary effort reported by Jay Lubin.

Following the subgroup sessions the plenary workshop considered the conclusions reached by the sub-groups. Predictably, there was considerable overlap between the discussions being reported on, and the following report is based on the reports of both sub-groups.

REPORT OF ANALYSIS SUB-GROUP

1. Analytical Issues

1.1 Basic Methods and Software

It was agreed that conditional logistic regression, as described by Breslow and Day in 1980, should be used to analyze the case-control data. There are several software packages that can be used to fit models of various forms based on this approach, such as EGRET, EPILOG and EPICURE; others may become available in the future.

1.2 Exclusion of individuals

Some investigators may wish to exclude from their study certain individuals who have received very specific exposures that may affect their lung cancer risk, e.g. exposure to asbestos or radiotherapy. Investigators should make clear whether they are excluding such individuals, and if so, provide relevant details. Since those excluded are likely to be small in number, it would not be necessary to place a term in the regression model to account for them.

1.3 Stratification Variables

(i) Age. Five-year age groups are likely to suffice. However, this could be checked by perhaps including a linear age term in the regression model as well as having age as a stratification variable.

(ii) Sex.

(iii) Race, where relevant. There should be caution about putting all non-whites into one group.

(iv) Area. In some studies area is not included as a design variable, whereas in others there is pair or frequency matching within areas. Given that adjusting for area can assist in removing urban/rural differences, it is important to stratify by area in the analysis.

(v) Interviewer. More information is required from investigators on how interviewers are used, so that any effect of interviewer may be taken into account in the analysis, either in the stratification or in the regression model.

(vi) Smoking. Some studies match on smoking. Whilst departures from a multiplicative model for the joint effect of radon and smoking can still be evaluated in this instance, the same cannot be done for departures from an

additive relationship. In studies that are unmatched for smoking, an alternative is to include smoking in the regression model (see below).

(vii) Other Stratifying Variables. It may be that in individual studies it will be necessary to stratify by a specific variable unique to the study; for example, by air pollution in the study in Shenyang, China (Blot et al, 1990). It is therefore not desirable to pool studies before individual investigators have performed analyses of their data.

1.4 Use of Hospital Controls

This is primarily a design issue. The use of controls who were initially suspected of having lung cancer may avoid a number of sources of potential sources of bias. It is, however, important to avoid including in the hospital control group patients whose current admission to hospital is for a disease strongly related to smoking. It is also desirable for patients with a wide variety of diseases to be included in the hospital control group.

1.5 Confounding Variables

Great care needs to be taken in making allowance for smoking. Breslow and Day (1980) had suggested using a regression model rather than stratification to adjust for smoking. In pooling data from different countries it needs to be recognized that the smoking risk varies between countries. One possibility therefore is to include terms in the regression model for interactions between smoking and study location.

Interviewer should be taken into account as a potential confounder in the regression model if not in the stratification. Among other variables for which adjustment may be made is some measure of social class or education.

1.6 Radon Assessment

Investigators are to be encouraged to look at data on the individuals' time spent in the dwellings. However it should be recognized that making questionnaires too detailed may lead to increasing inaccuracies in the replies. More information would be welcome on year-to-year variations in radon concentrations within a given house, so as to assist in estimating past exposures. Further methodological work is required on how to deal with measurement errors and missing data, not only for radon but also for confounders. Adjustment in the analysis for passive smoking and other factors that may influence the particle size distribution might help take account of any effect of particle size, as well as the effect of passive smoking per se.

2. Pooling

It would be desirable for individual investigators to perform the data management in preparation of pooling. Prior to pooling they should give thought to producing a common format for the data which gives more flexibility in the analysis.

In performing a test for trend in risk there are methodological issues that require investigation; for example, whether continuous or categorized forms of the radon measurements should be used in the analysis.

As stated earlier, differences in the risk from smoking between countries with different histories of smoking over time and between the genders need to be recognized when performing the analysis.

Sub-group analyses should be small in number and specified in advance.

There is a range of topics that can be addressed more easily in a parallel analysis, rather than in the analysis of individual studies. Such issues include:

(i) the form of exposure-response relationships based on cumulative and time-weighted radon indicies;

(ii) effect of exposure rate;

(iii) effect of living in a 'high' then a 'low' radon house, compared with vice versa;

(iv) effect of time-window for exposures;

(v) interaction of radon-related risk with attained age;

(vi) interaction of radon-related risk with sex;

(vii) interaction of radon-related risk with smoking;

(viii) interaction of radon-related risk with study-specific risk factors;

(ix) effect of cessation of smoking on the radon-related risk.

Detailed planning for the parallel analysis should be performed once it is known which studies are to be included, and the parallel analysis should not be performed before the individual investigators have analyzed their own data.

Plenary session

In the final plenary session the subgroup reports were presented and discussed. There was a consensus that the planning for pooling of the data that are in the process of being gathered should continue in view of the benefits to be expected from such pooling.

- It was understood that analysis of the individual studies should be carried out before a pooled analysis can be attempted, but at the same time communications and coordination between the various studies should be facilitated.

- It was agreed that such communication and coordination between studies could best be accomplished by setting up two nodal points, one in Europe and one in North America.

- Establishment and support of an electronic mail network to connect the current studies could be an important way of improving communications between the various studies.

The following Table is presented from the study data supplied by the participants to indicate the total number of cases and controls in each study, the approximate dates data collection started, and is expected to be completed, and the date the analysis of each of the individual studies is expected to be completed.

Investigator	Study Location	Cases	Controls	Start	Finish	Analysis
M.C.R. Alavanja	Missouri	600	1400	01/88	08/92	01/93
S.C. Darby	Devon, Cornwall	900	1800	01/88	01/93	01/95
J.L. Lyon	Utah, Idaho	750	1500	10/89	05/93	05/94
W.L. Nicholson	New Jersey	780	795	09/89	01/94	07/94
G. Pershagen	Stockholm County	210	400	01/83	01/89	07/91
G. Pershagen	Sweden	1500	3000	01/87	01/91	07/92
A. Poffijn	Ardennes, Eiffel	1200	3600	09/90	01/94	01/95
J.B. Schoenberg	New Jersey	480	442	01/85	01/91	01/92
H.G. Stockwell	Florida	300	450	01/88	01/92	07/92
J.A.J. Stolwijk	Connecticut	1000	1000	06/90	01/94	01/95
M. Tirmarche	Bretagne	600	1200	01/91	01/95	01/96
H.E. Wichmann	Germany	3000	3000	09/90	01/94	01/95

From this table it is apparent that extensive pooling of the data from current studies cannot be expected to be possible for several years.

ATTENDEES

Participants from ongoing studies:

Michael Alavanja
Epidemiology and Biostatistics Program
National Cancer Institute
Executive Plaza North, Room 543
Bethesda, MD 20892

Sarah C. Darby
Imperial Cancer Research Fund
Cancer Epidemiology and Clinical Trials Unit
University of Oxford
Gibson Building, The Radcliffe Infirmary
Oxford OX2 6HE

Richard McGregor
Bureau of Radiation and Medical Devices
Health and Welfare Canada
775 Brookfield Road
Ottawa, Ontario Canada K1A 1C1

William Nicholson
Department of Community Medicine
Mount Sinai Medical Center
1, Gustave L. Levy Place
New York, NY 10029-6574

Goran Pershagen
Institute of Environmental Medicine
Karolinska Institute
Box 60208
S 10401 Stockholm, Sweden

Andre Poffijn,
Laboratory of Physics
Proeftuin straat 86
B-9000 Gent, Belgium

Dale Sandler
Division of Biometry and Risk Assessment
National Institute of Environmental Health Sciences
P.O. Box 12233
Research Triangle Park, NC 27709

Janet Schoenberg
Chief, Cancer Epidemiology Program
New Jersey State Department of Health CN 369
Trenton, NJ 08625

Heather Stockwell
Department of Epidemiology and Biostatistics
College of Public Health
University of South Florida
13301 Bruce B. Downs Boulevard
Tampa, FL 33612

Jan A.J. Stolwijk
Department of Epidemiology and Public Health
Yale University School of Medicine
P.O. Box 3333
New Haven, CT 06510

Margot Tirmarche
CEA-ISPN/DPS/SEAPS
60-68 Avenue General Leclerc
BP 6 92265 Fontenay aux Roses CEDEX
France

Erich H. Wichmann
University of Wuppertal
FBSB 14
Gauss Strasse 20
D 5600 Wuppertal 1
Germany

Other participants:

Susan M. Conrath
Problem Assessment Branch
Radon Division
U.S. Environmental Protection Agency
401 M Street SW
Washington, DC 20460

Naomi Harley
New York University Medical Center
Department of Environmental Medicine
505 First Avenue
New York, NY 10016

Jay Lubin
National Cancer Institute
Biostatistics Branch
EPN/403
Rockville, MD 20892

Colin Muirhead
Assistant Director of Medicine
National Radiological Protection Board
Chilton, Didcot, Oxon
OX11 England ORO

Jerry Puskin
Environmental Protection Agency
401 M Street SW
Washington, DC 20460

Susan L. Rose
Radon Program Manager
Office of Health and Environmental Research
ER-72, E-222/GTN
19901 Germanown Road
U.S. Department of Energy
Germantown, MD 20874

Jonathan Samet
University of New Mexico School of Medicine
Cancer Research and Treatment Center
900 Camino de Salud NE
Albuquerque, NM 87131

Richard Sextro
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

Jaak Sinnaeve
Commission of the European Communities
Radiation Protection Programme
B-1049 Bruxelles
Belgium

Donna Spiegelman
Tufts University Sch. of Medicine
Department of Community Health
Division of Biometry
136 Harrison Avenue
Boston, MA 02111

Allan Tanner
12125 Captiva Court
Reston, VA 22091

APPENDIX A

Health Physics Vol. 60, No. 2 (February), pp. 223–227, 1991
Printed in the U.S.A.

0017-9078/91 $3.00 + .00

● *Note*

SUMMARY: INTERNATIONAL WORKSHOP ON RESIDENTIAL Rn EPIDEMIOLOGY

Jonathan M. Samet
New Mexico Tumor Registry, Cancer Center and the Department of Medicine, University of New Mexico, Albuquerque, NM 87131

and

Jan Stolwijk
Department of Epidemiology and Public Health, Yale University School of Medicine, P.O. Box 3333, New Haven, CT 06510

and

Susan L. Rose
Office of Health and Environmental Research, ER-73, Department of Energy, GTN, Washington, D.C. 20545

INTRODUCTION

RADON is an established cause of lung cancer for uranium miners and other ^{222}Rn-exposed underground miners. Radon is also present in the air of residences and other structures. The range of ^{222}Rn concentrations indoors is wide: In the United States, it ranges from well below the average of about 37 Bq m^{-3} (1 pCi L^{-1}) to levels of 3700 Bq m^{-3} (100 pCi L^{-1}) and higher. The higher concentrations reported in residences are comparable to concentrations in mines with well-documented excess lung cancer in the underground work force.

Although the presence of ^{222}Rn in indoor environments was well recognized in the scientific community during the 1970s, concern for the public-health risk posed by ^{222}Rn was not widely emphasized until the 1980s. In several European countries, programs for research, measurement, and control were implemented during the 1970s and early 1980s. In the United States, the almost accidental discovery in 1984 of a house in Pennsylvania with extremely high levels, and the subsequent finding of other homes throughout the country with unacceptably high levels of ^{222}Rn, made the media and public aware of the problem and prompted expanded measurement programs and initiatives for control.

Numerous epidemiological studies have addressed ^{222}Rn exposure and lung cancer in underground miners; many of these studies provide quantitative risk estimates, and several offer evidence on the combined effect of smoking and ^{222}Rn exposure. To date, the risks of indoor ^{222}Rn have been estimated primarily by extrapolating from the results of the studies of miners to the indoor environment. This extrapolation is subject to uncertainties reflecting differences between the mining and indoor environments, differing biological characteristics of the miners (males exposed during adulthood, with heavy particulate exposure) and the general population (males and females exposed at all ages), and the limited information on the lifetime expression of the excess lung cancer risk associated with ^{222}Rn and its interaction with risks due to smoking.

Epidemiological studies conducted in the general population represent an alternative and potentially more effective and persuasive approach for assessing the lung cancer risk associated with indoor ^{222}Rn. Several preliminary studies of this type have been reported and many others are planned or underway in the United States and abroad. As in many other areas of environmental epidemiology, these studies are subject to constraints posed by feasibility; by difficult methodological problems related to exposure assessment, confounding, and statistical power; and by available resources.

In early 1988 it was still unclear how many of these studies were ongoing, and more importantly, how comparable the study designs were if combined analyses were to be undertaken in the future. The U.S. Department of Energy, the lead U.S. agency for ^{222}Rn-related basic research, supported Dr. John S. Neuberger of the University of Kansas to identify epidemiological studies of indoor ^{222}Rn throughout the world. The Department of Energy

(*Manuscript received* 14 *May* 1990; *revised manuscript received* 5 *July* 1990, *accepted* 1 *August* 1990)

considered the cataloging of these studies to be important and planned to convene a workshop of all active investigators.

On 24–26 July 1989, with the sponsorship of the Department of Energy and the Commission of the European Communities, a meeting was held in Alexandria, VA that brought together 18 participating scientists and additional observers, representing 10 countries (see Appendix). The studies represented by the participating scientists were located in Belgium, Canada, the Federal Republic of Germany, Finland, France, Norway, Sweden, the Peoples' Republic of China, the United Kingdom, and the United States. The observers were from the Department of Energy, the Commission of the European Communities, the U.S. Environmental Protection Agency, the U.S. Public Health Service, the International Atomic Energy Agency, the International Agency for Research on Cancer, and the health departments of the states of California, Minnesota, and New Jersey.

The agenda was designed to facilitate awareness among the investigators of the work in progress, to review common methodological problems, and to address the potential and desirability of pooling the data from the individual investigations. The workshop began with a review of the individual projects. Subsequently, four working groups were formed to review specific areas: Exposure Assessment, Study Design, Analysis and Meta-analysis, and Policy. Recommendations from the four working groups were then reviewed by all workshop participants.

The findings of each group, as well as the overall conclusions, are presented in this article. The full proceedings of the workshop contain a more-complete description and summaries of the individual studies. The complete proceedings can be obtained from the National Technical Information Service (International Workshop on Residential Radon Epidemiology, CONF-8907178; National Technical Information Service, U.S. Department of Commerce, Springfield, VA 22161).

WORKSHOP SUMMARY

Exposure assessment

This working group considered four separate dimensions of exposure assessment: factors affecting exposure and exposure-dose relations; measurement techniques, protocols, and quality assurance; measurements in prior homes and occupancy; and estimation of past concentrations when measurements cannot be made. With regard to the first issue, the working group recommended that measurement of ^{222}Rn alone, without measuring concentrations of progeny or equilibrium, was appropriate for epidemiological studies. The group also concluded that particle-size distributions and particle concentrations were of little utility in epidemiological studies of indoor ^{222}Rn.

Filtered etched-track detectors were recommended for epidemiological studies, with placement, preferably for a 1-y period, in the living space. Short-term techniques were considered suitable only to provide data for houses for which etched-track data cannot be obtained. A well-designed quality control and assurance program should be part of any study. The development of a common format for reporting exposure over time was also recommended.

Locating past homes and making measurements in them represents a difficult problem. Major changes in the characteristics of prior houses should be identified, especially changes made in the substructure or for the purpose of ^{222}Rn mitigation. The working group also encouraged the collection of information on time-activity patterns of the subjects.

Study design

The conduct of epidemiological studies of ^{222}Rn and lung cancer raises many difficult methodological problems; some of these problems are general and inherent in environmental epidemiology, whereas others are specific to investigating ^{222}Rn and lung cancer. This group focused on the latter issues, while recognizing that the general issues must also be addressed in all studies of ^{222}Rn and lung cancer.

The group proposed that new epidemiological studies of indoor ^{222}Rn and lung cancer should be designed to address the exposure-effect relationship and not solely to test whether ^{222}Rn exposure increases lung cancer risk. A study should be designed to provide a sufficiently narrow risk estimate so that the risk estimate is useful for the purpose of quantitative risk assessment. Overall, the designs of the various epidemiological studies should collectively cover the most important potential modifiers of the effect of ^{222}Rn, with particular emphasis on tobacco smoking.

Characteristics of the study area and population should be considered in implementing a study. Populations with low mobility may be most informative, and especially studies aimed at childhood exposure are best carried out in populations with highly stable residence locations. The study area should also offer a broad range of exposure. Other features to be considered include the lung cancer incidence in the area or the population, and the availability of a mechanism for locating and ascertaining cases.

In considering alternative study designs, ecological or analytical, the group recommended that analytical studies be undertaken and that ecological studies should not be planned unless warranted by a special situation or unique opportunities. Of the two analytical designs, the case-control study is the preferred approach. Special circumstances may provide justification for undertaking a cohort study.

In epidemiological studies on domestic ^{222}Rn exposure and lung cancer, a number of other factors have to be considered as effect modifiers and/or confounders. Demographic information, smoking status, passive smoking exposure, and occupational risks need to be determined in as quantitative a manner as possible, if confounding is to be controlled and interaction properly as-

sessed. As a minimum, the data on smoking should include information on ages of starting and stopping, daily consumption, and types of tobacco products smoked.

Careful validation of the diagnosis of lung cancer was encouraged, preferably using histological confirmation. For studies in which histological type is determined, procedures for review of histopathological materials should be implemented. Issues related to selection of controls were also addressed.

The working group commented that unique ethical issues may be encountered in studies of ^{222}Rn and lung cancer. Owners and/or occupants of a house with high levels must be informed and advised. Confidentiality regarding ^{222}Rn levels should also be respected.

Analysis and meta-analysis

This working group reviewed sample size needs for studies of ^{222}Rn and lung cancer. On the basis of calculations presented at the workshop, it was recognized that many of the on-going and planned studies had samples too small, by themselves, to achieve the objective of characterizing the exposure-effect relationship. Misclassification of exposure was also identified as a potentially significant problem in epidemiological studies. The difficulty of analyzing the time-dependent relationship between ^{222}Rn and lung cancer risk was acknowledged and additional research in this area was encouraged.

The group recommended that consideration should be given to pooling the data from the various studies. The diversity of protocols and populations among the studies was seen as a strength in this regard. Specific aspects of pooling were discussed. Initial pooled analysis might be based on published data that provide point estimates of the excess relative risk per unit exposure. Further pooled analyses should be carried out by a panel representing as many teams as possible that carried out the original studies. Alternative approaches for pooling were considered. In studies being planned, it was recommended that the protocol and study design be so structured as to facilitate pooling. Minimum data requirements were reviewed. The working group also described a Bayesian approach to meta-analysis.

Policy

The group's discussion reflected a strong sentiment that the scientific evidence on lung cancer and ^{222}Rn should guide the evolution of public policy, along with other considerations such as feasibility and costs of mitigation. The working group identified nine questions for which policy makers are seeking answers:

1. How serious is the lung cancer risk associated with indoor ^{222}Rn? *In homes? In schools? In workplaces?*
2. What are the risks at different levels of exposure?
3. How is the risk of lung cancer modified by smoking? By other factors?
4. Are children at greater risk?
5. What is the optimum risk management strategy, given our present understanding of the individual and population risks, and the feasibility of mitigation and costs?
6. How can ^{222}Rn risks be communicated with perspective?
7. Who is responsible for controlling ^{222}Rn?
8. Is it a government's role to overcome public apathy?
9. (For the U.S.) How can the federal government assist the states?

Epidemiological studies can directly address only the first four questions, but epidemiological data are also relevant for questions 5 and 6. Moreover, policy makers should be able to proceed with a clearer purpose on the last three items as scientific understanding of ^{222}Rn and lung cancer deepens.

The studies in progress can be expected to provide new information relevant to questions 1 through 4, but uncertainties are likely to remain. Certain questions cannot be answered, however. For example, the epidemiological evidence is unlikely to meet the demands of legal criteria for causality. Policy makers should not use the inability of epidemiologists to identify specific individuals with ^{222}Rn-induced lung cancer as a basis for minimizing or dismissing the problem. Pooling of past and present studies was considered an appropriate strategy for providing more precise answers to the policy maker's questions.

SUMMARY

The participants in the workshop agreed with the need for communication between the investigators to facilitate the individual studies and to foster the development of a plan for pooling. The participants also agreed that a future workshop should be devoted to methods for pooling and even analysis of published data. It was also recommended that a mechanism be set up to facilitate and encourage pooling of data from different studies now ongoing or planned. This mechanism should address the development of procedures for pooling and protocol requirements. The development of this mechanism is in progress.

APPENDIX

Meeting participants, with their respective working groups indicated in parentheses.

J. Ahmed (Exposure Assessment), International Atomic Energy Agency, Wagramerstrasse, 5, P.O. Box 100, 1400 Vienna, Austria

Michael Alavanja (Analysis), Epidemiology and Biostatistics Program, National Cancer Institute, Executive Plaza North, Room 543, Bethesda, MD 20892, United States

Olav Axelson (Study Design), Department of Occupational Medicine, University Hospital, 581 85 Linkoping, Sweden

Susan M. Conrath (Study Design), Problem Assessment Branch, Radon Division, U.S. Environmental Protection Agency, 401 M Street SW, Washington, D.C. 20460, United States

Kathleen Conway (Policy), U.S. Environmental Protection Agency, Science Advisory Board (A101F), 499 S. Capitol Street SW, Room 508, Washington, D.C. 20460, United States

Sarah Darby (Analysis), Imperial Cancer Research Fund, Oxford University, Cancer Epidemiology and Clinical Trials Unit, Gibson Building, The Radcliffe Infirmary, Oxford OX2 6HE, England

Christopher Dodge (Policy), Library of Congress, SPRD/CRS, LM 413, Washington, D.C. 20540, United States

Jacques Esteve (Policy), International Agency for Research on Cancer, 150 Cours Albert-Thomas, 69372 Lyon Cedex D8, France

Elizabeth Fontham (Study Design), Department of Pathology, Louisiana State University Medical Center, 1901 Perdido Street, New Orleans, LA 70112, United States

Bill Forster, Office of Health and Environmental Research, U.S. Department of Energy, ER-75 E-239/GTN, Germantown, MD 20545, United States

Naomi Harley (Exposure Assessment), New York University Medical Center, Department of Environmental Medicine, 550 First Avenue, New York, NY 10016, United States

David Hoel, Division of Biometry and Risk Assessment, National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, NC 27709, United States

Sara Hospodor (Policy), National Association of Realtors, 777 14th Street NW, Washington, D.C. 20005, United States

Zdenek Hrubec (Exposure Assessment), Radiation Epidemiology Branch, National Cancer Institute, Executive Plaza, North—Suite 408, 6130 Executive Blvd., North Bethesda, MD 20852, United States

Judith Klotz (Policy), Cancer Epidemiology Program, New Jersey Department of Health, CN 369, Trenton, NJ 08625, United States

Ernest Letourneau (Policy), Bureau of Radiation and Medical Devices, Health and Welfare Canada, 775 Brookfield Road, Ottawa, Ontario K1A 1C1, Canada

Kai-Shen Liu (Policy), California Department of Health Services, Air and Industrial Hygiene Laboratory, 2151 Berkeley Way, Room 334, Berkeley, CA 94704, United States

Jay Lubin, National Cancer Institute, Biostatistics Branch, EPN/403, Rockville, MD 20892, United States

Joseph Lyon, Department of Family and Preventive Medicine, University of Utah, 50 North Medical Drive, Salt Lake City, UT 84132, United States

Judith Mahaffey (Exposure Assessment), Pacific Northwest Laboratory, P.O. Box 999, MSIN K1-85, Richland, WA 99352, United States

William Mills (Policy), Oak Ridge Associated Universities, 1019 19th Street NW, Suite 700, Washington, D.C. 20036, United States

Colin Muirhead (Exposure Assessment), Assistant Director of Medicine, National Radiological Protection Board, Chilton, Didcot, Oxon, OX11 ORQ, England

Anthony Nero, Jr. (Exposure Assessment), Indoor Environmental Program, Applied Science Division, Lawrence Berkeley Laboratory, University of California, Berkeley, CA 94720, United States

John Neuberger, Department of Preventive Medicine, University of Kansas School of Medicine, Rainbow Blvd. at 39th, Kansas City, KS 66103, United States

Mark Nowak (Policy), National Association of Homebuilders, 15 & M Streets NW, Washington, D.C. 20005, United States

Laura Oatman (Policy), Indoor Air Quality Unit, Minnesota Department of Health, 717 SE Delaware Street, P.O. Box 9441, Minneapolis, MN 55440, United States

Goran Pershagen (Study Design), Institute of Environmental Medicine, Karolinska Institute, Box 60208, S 10401, Stockholm, Sweden

Andre Poffyn (Exposure Assessment), Lab of Physics, Proeftuinstraat 86, B-9000 Gent, Belgium

Jerry Puskin (Analysis), Office of Radiation Programs, U.S. Environmental Protection Agency, 401 M Street SW, Washington, D.C. 20460, United States

Tom Rohan (Analysis), Epidemiology Unit, National Cancer Institute of Canada, University of Toronto, 12 Queen's Park Cres. West, Toronto, Ontario M5S 1A8, Canada

Susan Rose (Policy), Office of Health and Environmental Research, U.S. Department of Energy, ER-73, E-222/GTN, 19901 Germantown Road, Germantown, MD 20874, United States

Eeva Ruosteenoja (Analysis), Finnish Centre for Radiation and Nuclear Safety, P.O. Box 268, SF-00101 Helsinki, Finland

Jonathan Samet (Study Design), University of New Mexico School of Medicine, Cancer Research and Treatment Center, 908 Camino de Salud NE, Albuquerque, NM 87131, United States

Janet Schoenberg (Analysis), Chief, Cancer Epidemiology Program, New Jersey State Department of Health, CN 369, Trenton, NJ 08625, United States

Stuart Shalat (Analysis), Yale-Connecticut Radon Study, Department of Epidemiology and Public Health, Yale University School of Medicine, 530 Whitfield Street, Suite Two, Guilford, CT 06437, United States

Jaak Sinnaeve (Policy), Commission of the European Communities, Rue de La Loi 200, Radiation Protection Programme, (XII-D-3/ARTS 2-47), B-1049 Bruxelles, Belgium

James Stebbings (Exposure Assessment), Environmental Health Section, Division of Biological and Medical Research, BEM, 203J, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439, United States

F. Steinhausler (Study Design), University of Salzburg, Division of Biophysics, Hellbrunnerstr 34, A-5020 Salzburg, Austria

Heather Stockwell (Study Design), Department of Epidemiology and Biostatistics, College of Public Health, University of South Florida—MHH104, 13301 Bruce B. Downs Boulevard, Tampa, FL 33612, United States

Jan Stolwijk (Analysis), Department of Epidemiology and Public Health, Yale University School of Medicine, P.O. Box 333, New Haven, CT 06510, United States

Terje Strand (Study Design), National Institute of Radiation Hygiene, P.O. Box 55, N-1345 Osteras, Norway

John Talbott (Exposure Assessment), Office of Conservation and Renewable Energy, MS-CE132, U.S. Department of Energy, 1000 Independence Avenue, Washington, D.C. 20585, United States

Kevin Teichman (Policy), Office of Technology Transfer and Regulatory Support, U.S. Environmental Protection Agency, 401 M Street SW, Washington, D.C. 20460, United States

Margot Tirmarche (Policy), CEA-IPSN/DPS/SEAPS, 60-68 Avenue General Leclerc, BP 6 92265 Fontenay aux roses cedex, France, 46547194

Richard Toohey (Policy), Argonne National Laboratory, Environmental Safety and Health Department, 9700 South Case Avenue—Bldg 201, Argonne, IL 60439, United States

Michel Vanhoorne (Study Design), Department of Hygiene and Social Medicine, University Hospital, De Pinteloon, 185, 9000 Gent, Belgium

Walt Warnick (Policy), U.S. Department of Energy, Office of Health and Environmental Research, ER-32, F-327/GTN, Germantown, MD 20874, United States

Claire Weinberg (Exposure Assessment), Division of Biometry and Risk Assessment, National Institute of Environmental Health, P.O. Box 12233, Research Triangle Park, NC 27709, United States

Erich Wichmann (Analysis), University of Wuppertal, FB 14, Gauss Street No. 20, D 5600, Wuppertal 1, West Germany

Investigations in Progress

Country: Belgium
P.I.(s): Andre Poffijn
Collaborator(s): R. Mak, M. Vanhoorne, P. Weynants
Study Type: Case-control, 100 cases and 200 controls

Country: Canada
P.I.(s): E. G. Letourneau
Collaborator(s): N. W. Choi, Roger S. Eaton, R. McGregor, D. C. Wiggle
Study Type: Case-control, 700 cases and 700 controls

Country: Canada
P.I.(s): G. R. Howe
Collaborator(s): Health and Welfare Canada
Study Type: Case-control, 300 cases and 1500 controls

Country: Finland
P.I.(s): E. Rousteenoja
Collaborator(s): M. Hakama, I. Makelainen, R. Rytomaa
Study Type: Nested case-control, 291 cases and 450 population controls

Country: France
P.I.(s): Margot Tirmarche
Collaborator(s): National Institute of Health and Medical Research (INSERM), Univ. of Brest, Laboratory of Aerosol Physics and Atmospheric Radioactivity
Study Type: Case-control, 600 cases and 1200 controls

Country: Federal Republic of Germany
P.I.(s): H. Erich Wichmann
Study Type: Case-control, 3200 cases and 3200 controls

Country: Sweden
P.I.(s): Goran Pershagen
Collaborator(s): Zdenek Hrubec
Study Type: Case-control, 210 cases and 209 population controls and 191 controls with suspected but unconfirmed lung cancer

Country: United Kingdom
P.I.(s): Sara C. Darby, P. B. S. Silcocks, Sir Richard Doll
Collaborator(s): National Radiological Protection Board
Study Type: Case-control, 600 cases and 1200 controls

Country: Connecticut, United States
P.I.(s): Stuart L. Shalat
Collaborator(s): Jan Stolwijk, Theodore Holford, Brian Leaderer
Study Type: Case-control, 1000 cases and 1000 controls

Country: New Jersey, United States
P.I.(s): Janet B. Schoenberg
Collaborator(s): Judith B. Klotz, Gerald P. Nicholls, Zdenek Hrubec, Thomas J. Mason
Study Type: Case-control, 433 cases and 402 controls

Country: Louisiana, United States
P.I.(s): Elizabeth T. H. Fontham
Collaborator(s): Pelayo Correa, Raymond Greenberg, Peggy Reynolds, Anna Wu, Patricia Buffler
Study Type: Case-control, 200 cases and 650 controls

Country: Missouri, United States
P.I.(s): Michael C. R. Alavanja
Collaborator(s): Ross Brownson, Zdenek Hrubec, Sheila Hoar Zahm
Study Type: Case-control, 500 cases and 1000 controls

Country: Pennsylvania, United States
P.I.(s): Dr. James H. Stebbings
Study Type: Case-control, 500 cases and 500 controls

Country: Utah and Southern Idaho, United States
P.I.(s): Joseph L. Lyon
Collaborator(s): U.S. National Institute of Environmental Health Sciences
Study Type: Case-control, 775 cases and 1125 controls

APPENDIX B

Health Physics Vol. 59, No. 6 (December) pp. 807–817, 1990
Printed in the U.S.A.

0017-9078/90 $3.00 + .00

● *Paper*

DESIGN ISSUES IN EPIDEMIOLOGIC STUDIES OF INDOOR EXPOSURE TO Rn AND RISK OF LUNG CANCER

Jay H. Lubin
National Cancer Institute, Epidemiology and Biostatistics Program, Epidemiologic Methods Section, 6130 Executive Blvd., Room 403, Rockville, MD 20892

and

Jonathan M. Samet
Department of Medicine and New Mexico Tumor Registry, Cancer Center, University of New Mexico Medical Center, Albuquerque, NM 87131

and

Clarice Weinberg
National Institute of Environmental Health Sciences, Division of Biometry and Risk Assessment, P.O.B. 12233, Research Triangle Park, NC 27709

Abstract—**Recent data on indoor air quality have indicated that Rn (^{222}Rn) and its decay products are frequently present in domestic environments. Since studies of Rn-exposed miners have established that Rn decay products are a lung carcinogen, their presence in indoor air raises concerns about an increase in lung cancer risk for the general population. To directly evaluate lung cancer risk from domestic exposure to Rn and its decay products, as well as to evaluate risk assessments derived from studies of Rn-exposed underground miners, several epidemiologic studies of indoor Rn exposure have been initiated or are planned. This paper calculates sample sizes required for a hypothetical case-control study to address several important hypotheses and shows the impact of several difficult problems associated with estimating a subject's Rn exposure. We consider the effects of subject mobility, choice of the exposure response trend which is used to characterize an alternative hypothesis, and errors in the estimation of exposure. Imprecise estimation of Rn exposure arises from errors in the measurement device, exposure to Rn decay products from sources outside the home, inability to measure exposures over time in current as well as previous residences, and the unknown relationship between measured concentration and lung dose of α energy from the decay of Rn and its progeny. These methodological problems can result in large discrepancies between computed and actual study power. Failure to anticipate these problems in the design of a study can result in inaccurate estimates of power. We conclude that case-control studies of indoor Rn and lung cancer may require substantial numbers of subjects in order to address the many questions of importance that burden current risk assessments with uncertainty. We suggest pooling data from studies with the largest numbers of cases and with the most precise estimates of Rn exposure as the best approach for meeting present research needs.**

INTRODUCTION

AS DATA on indoor air quality were collected during the 1980s, it became apparent that Rn and its decay products are frequently present in indoor environments and that concentrations may reach very high levels in some homes. The presence of an established carcinogen in indoor air raised concern that exposure to Rn decay products may increase risk of lung cancer for the general population, including both smokers and nonsmokers.

As a basis for determining the acceptability of the risks associated with Rn in indoor air, risk assessment methods have been used (ICRP 1987; Lubin and Boice 1989; NAS 1988; NCRP 1984; U.S. EPA 1986); the most widely employed risk projection models extrapolate from studies of the risk of lung cancer in Rn-exposed underground miners to the general population. Extrapolation from studies of miners has been necessary because data relating lung cancer risk to domestic exposure to Rn are as yet insufficient.

Risk projections from miner cohorts to the general population are subject to uncertainty for several reasons. Exposure concentrations are generally higher in mines; the subjects have been only males; the age at which underground miners began employment covered a restricted

(*Manuscript received* 14 *December* 1989; *revised manuscript received* 3 *May* 1990, *accepted* 14 *May* 1990)

range and did not extend into childhood, except for one study in China of underground tin miners (Lubin et al. 1990; Qiao et al. 1989); and follow-up still does not extend across the full lifetimes of the cohort members. In addition, miner-based models are themselves subject to various sources of uncertainties. Moreover, it is unclear how best to apply miner-based models since, because of different environmental conditions and patterns of activity, the carcinogenic effect of Rn exposure may be different in homes and mines. Only a few of the principal cohort studies include the information on cigarette smoking needed to assess the combined effect of smoking and Rn exposure; furthermore, in those cohorts where smoking status is available, few workers were nonsmokers.

To reduce uncertainty concerning the risk of Rn in indoor air, ecologic and etiologic studies on lung cancer and exposure to Rn in indoor air have been undertaken (Borak and Johnson 1988; New Jersey Health Department 1989; Samet 1989). In ecologic studies, lung cancer rates for geographic units have been correlated with factors suggestive of exposure to Rn, generally based on geological morphology, surface radioactivity, or limited numbers of measurements. These studies have provided only modest evidence of an association of domestic exposure with lung cancer risk. This study design, however, has well-characterized limitations (Morgenstern 1982; Piantadosi et al. 1988); ecological correlations may be substantially biased in estimating the effect at the level of the individual; an estimate of current Rn exposure may have little relevance for lung cancer risk; and confounding by cigarette smoking or a correlate may not be adequately controlled. Incomplete control of confounding due to smoking could easily obscure the effect of Rn. For example, suppose smoking increases lung cancer risk 10-fold and that 30% of "non-Rn-exposed" persons smoke, whereas 36% of the "Rn-exposed" persons smoke. Using formula 2 in Gail et al. (1988), a 15% bias is introduced if cigarette smoking is not considered.

Case-control and cohort designs can potentially overcome some of these limitations. The case-control design has been the principal etiological approach for studying indoor Rn and lung cancer, since the cohort study design, involving extended follow-up of large groups of subjects, has not generally been feasible because of the inherent difficulties in maintaining contact and monitoring exposure. Moreover, most case-control and cohort studies have been small in size and the findings too imprecise to meet risk assessment needs (Samet 1989).

In our view, the existing evidence has established that exposure to Rn decay products can cause lung cancer in humans (ICRP 1987; NAS 1988; NCRP 1984). Therefore, epidemiologic studies of domestic Rn exposure should not focus solely on reestablishing the association of Rn with lung cancer. Studies should also address issues of current scientific and public health importance, such as the quantitative magnitude of the risk of lung cancer from exposure to domestic sources of Rn, since effects might plausibly differ in homes and in mines (James 1988; NAS 1988; Samet 1989); the role of gender in determining risk; the effects of exposure rate, age at exposure, temporal patterns of exposure and time since exposure; and the combined effect of cigarette smoking and Rn and the effect of Rn in nonsmokers.

Although case-control studies can theoretically address the principal questions related to indoor Rn, we discuss in this paper some formidable methodological problems that make them difficult to carry out and that may limit the informativeness of the resulting data. We assess the impact of several problems on the required size of case-control studies of the association of lung cancer and domestic Rn exposure.

Rn EXPOSURE AND LUNG CANCER RISK

We begin by describing a model for the relationship between Rn exposure and lung cancer risk, which was developed from studies of Rn-exposed miners. We use this model to evaluate the potential of case-control studies to address research needs related to domestic Rn exposure. Although the studies of miners have limitations, they provide the best data at the present time for developing models of lung cancer risk. Under common residential occupancy patterns and Rn concentration assumptions, mean lifetime exposure for the general population is estimated as 12–20 WLM,* while mean cumulative exposure in the workplace for the populations of exposed miners considered by the National Academy of Sciences' Committee on the Biological Effects of Ionizing Radiation (referred to as the BEIR IV Committee) (NAS 1988) ranged from 22 to 509 WLM. Although Rn concentrations in homes are generally much lower than in mines, concentrations in some homes can yield cumulative exposures in the range received by some underground miners.

To develop a model for lung cancer risk, let W represent cumulative exposure to Rn decay products in WLM. For simplicity, we assume that the relative risk over a limited age interval can be modeled as linear in cumulative exposure. Because lung cancer is relatively rare, the odds ratio, $R(W)$, is also approximately linear, so that we may write:

$$R(W) = 1 + \beta \times W, \qquad (1)$$

where β is the increase in the excess relative risk per unit increase in WLM. The probability of disease during the

* Working Level Months is the product of time, in units of working month, taken to be 170 h, and Working Levels, a measure of radiation energy. One Working Level equals any combination of Rn progeny in 1 L of air, which results in the emission of 130,000 MeV of potential energy from α particles. Radon levels in the home are generally measured in units of Bq m^{-3} (or pCi L^{-1}), a count of the number of atomic transformations per second per cubic meter (or liter). Based on Rn at 50% equilibrium with its decay products and on persons resident in their homes 70% of the time, yearly exposure to 37 Bq m^{-3} in the home is approximately equal to 0.2 WLM y^{-1}.

age interval under study for exposure W, denoted $P(D = 1|W)$, can be expressed as

$$P(D = 1|W) = \frac{e^{\alpha}R(W)}{1 + e^{\alpha}R(W)},$$

with $P(D = 0|W) = 1 - P(D = 1|W)$, where $D = 1$ denotes the occurrence of lung cancer and $D = 0$ denotes absence of lung cancer. The quantity $\exp(\alpha)$ is the relative odds in the absence of exposure. For model 1, the BEIR IV Report estimated $\beta = 0.015$, that is, a 1.5% increase in the excess relative risk per WLM (NAS 1988).

Because cigarette smoking is the dominant cause of lung cancer in the United States, the risk from exposure to Rn decay products must be addressed for both smokers and nonsmokers. Evidence on the interaction between cigarette smoking and exposure to Rn decay products is limited but suggests that there is synergy. Recent analyses of data from the study of Colorado Plateau miners reject an additive relative risk model and indicate that the interaction is best described by a submultiplicative model, although the data are statistically consistent with a multiplicative model (NAS 1988). Extending the disease probability model given in expression 1, we replace exposure W by the bivariate exposure vector (W, S), where S is mean number of cigarettes smoked per day during the time the subject was a smoker. The relative risk now has the general form $R(W, S)$. Using an approach suggested by Thomas (1981), we define the joint odds ratio model as a composite of multiplicative and additive components with an indexing parameter, θ, viz.,

$$R(W, S) = [R(W) \times R(S)]^{\theta}[R(W) + R(S) - 1]^{1-\theta}, \quad (2)$$

where $R(W)$ and $R(S)$ are the odds ratio patterns for each exposure. In the joint model, $\theta = 1$ specifies a multiplicative relationship, while $\theta = 0$ specifies an additive relationship. Intermediate models arise if $0 < \theta < 1$ and more extreme models if $\theta < 0$ or $\theta > 1$. Analysis of miner data suggests that a linear relationship for the odds ratio is appropriate for W and for S, that is, $R(W) = 1 + \beta \times W$ and $R(S) = 1 + \gamma \times S$, where β and γ are the excess risk parameters for W and S, respectively (Lubin and Gaffey 1987).

In the context of models 1 and 2, three hypotheses can be formulated (Table 1). For the purposes of public policy and motivating the public to measure and mitigate indoor Rn, many of the initial etiologic studies of indoor Rn and lung cancer have been designed to test that domestic Rn exposure does not cause lung cancer (hypothesis 1). Other possible design objectives include the appropriateness of applying the exposure-response relationship for lung cancer risk derived from cohorts of miners to domestic exposure to Rn (hypothesis 2) and the interaction of Rn exposure with cigarette smoking (hypothesis 3).

Table 1. Hypotheses of potential interest in an epidemiologic investigation of indoor Rn and lung cancer.

1. No association: Rn exposure in the indoor environment does not cause lung cancer, i.e., $\beta = 0$ in Model 1.
2. Quantitative risk: The effect of Rn exposure in the indoor environment is lower by a factor of one-half than that in the mining environment, e.g., for $\beta = 0.015$ under the null hypothesis and $\beta = 0.0075$ under the alternative.
3. Interaction: The interaction between cigarette smoking and Rn exposure takes a particular form under the null hypothesis relative to a fixed alternative; e.g., the interaction is additive under the null hypothesis ($\theta = 0$ in Model 2) and multiplicative under the alternative ($\theta = 1$ in Model 2).

CASE-CONTROL STUDIES OF Rn AND LUNG CANCER

Overview

Conceptually, the design of a case-control study of indoor Rn and lung cancer is straightforward. Cases of lung cancer in a region of potential interest are identified from a regional tumor registry, death records, hospital discharge or pathology logs, or other sources. Controls are selected from the same region with some form of matching on age, gender, and possibly other factors. An interview is conducted with the cases or their next-of-kin concerning smoking habits, occupation, and other risk factors. The current and former residences are also identified. The same information is obtained for controls. An attempt is then made to measure Rn concentrations in current and past residences, using a short-term or long-term test protocol. Cumulative exposure to Rn can be estimated by integrating the measured Rn concentrations over time.

Case-control studies of lung cancer and domestic exposure to Rn pose unique and difficult problems related to exposure assessment (Table 2). To estimate exposure, Rn concentration must be measured in one or more homes over intervals of up to 1 y. Present concentrations may not adequately reflect past concentrations, and data may be missing because residences are unavailable for measurement. Radon concentrations must then be used to estimate exposure of an individual. Our analyses specifically address the implications of problems in assessing Rn exposure and in selecting the proper design assumptions. Other limitations inherent to case-control studies of lung cancer—for example, misclassification of disease and use of data from interviews with surrogates—have been discussed elsewhere (Garfinkel et al. 1985; Gordis 1982; Lerchen and Samet 1986; McFarlane et al. 1986, 1987; Pickle et al. 1983).

Estimation of exposure to Rn

Exposure to residential Rn is based on estimates of the annual concentration of Rn in the living area for each residence from birth to age at interview and the fraction of each year the subject was resident in the home. If Rn exposure is to be converted to exposure to Rn decay products, an equilibrium value of Rn with its decay products must be assumed. Note that estimated exposure to Rn

Table 2. Potential limiting factors of case-control studies of indoor Rn and lung cancer.

Sources of errors in estimation of Rn exposure
Measurements of Rn concentration
Characterization of historical concentrations
Residential history
Other sources of Rn exposure
Equilibrium fraction for Rn and its decay products
Relationship of Rn exposure to lung dose
Errors in estimation of tobacco use and other potential confounders
Missing data due to nonresponse
To interview
To Rn measurements
From use of surrogate responders
Misclassification of disease
Inappropriate design assumptions
Incorrect specification of the true value of β
Incorrect specification of the true exposure distribution
Failure to consider the consequences of residential mobility
Failure to consider the effects of random error in estimation of Rn exposure

(in Bq m^{-3}) or to Rn decay products (in WLM) is only a crude surrogate for the more biologically meaningful quantity of lung dose of α energy.

The accuracy of estimates of cumulative exposure for an individual depends on the accuracy of the concentration measurements and of the residential and occupancy history, and on the determination of the equilibrium fraction, if exposure to Rn decay products is estimated.

The measured Rn concentration depends on the measurement device and on placement of the device within the home. Several types of detectors are currently available for either long-term measurements (3–12 mo) or short-term measurements (2–7 d). Short-term devices, while attractive from an operational point of view, have the disadvantage of greater variation due to seasonal, as well as diurnal, factors (Cohen and Gromicko 1988; Hess et al. 1983, 1985; Nero et al. 1986; Ronca-Battista and Magno 1988) and should generally be avoided in case-control studies.

The placement of the measurement device is an important determinant of Rn levels. Studies suggest that there may be as much as a two- to fivefold difference in Rn concentrations between the basement and upper floors of a single-family dwelling (Alter and Oswald 1987; Cohen and Gromicko 1988; Crameri et al. 1989; Fleischer et al. 1983; George and Breslin 1980; Jönsson 1988). The EPA recommendation, that devices be placed in the lowest livable area of the home, was guided by the need to screen homes for maximum Rn levels and thereby reduce false negative rates (Puskin and Nelson 1989). In epidemiologic studies, accurate characterization of exposure suggests placement in "common living areas," such as a family room, kitchen, and bedrooms. Unfortunately, even this procedure may be subject to substantial variability, as micro-environmental conditions within a room may affect measured values. In one case-control study in which pairs of year-long Rn measurements were made, most of which were located within a single room, over 20% of the pairs of Rn measurements were found to vary by more than a factor of 2.[†]

Another source of exposure uncertainty arises from measurement error of the detector itself. Calibration error may be on the order of 10% but possibly as high as 20% (Bierma et al. 1989). For α-track devices, counting errors are another source of inaccuracy and depend on the Rn concentration (Bierma et al. 1989).

In case-control studies, exposures must be retrospectively estimated by first locating past residences and then measuring exposure concentrations. It must be assumed that measured Rn concentrations represent exposures prevailing during the years of occupancy of the subject. *Ad hoc* adjustments may have to be made to account for reported changes in the characteristics of the dwelling, such as weatherization.

To calculate cumulative Rn exposure, time spent in the home, derived from duration and percent occupancy time, must also be estimated. Total duration of occupancy can be assessed, probably with adequate accuracy, by taking a lifetime residency history. Time spent at home varies over a lifetime, depending on time spent at school, at work, and on other activities; recall may be a problem, especially for people who are elderly and sick. Moreover, attention has been generally restricted to home exposures, although additional exposure to Rn may occur outside the home. However, available data suggest that, except under unusual circumstances, exposures outside the home likely contribute minimally to total cumulative exposure, particularly if employment is out of doors or in a multilevel building. (The effects of estimating exposures only for the recent past, 5–30 y, are considered in the "Discussion.")

Missing data on exposure

In a case-control study of domestic Rn exposure and lung cancer, complete information needed to estimate exposure will inevitably be lacking for some subjects. Gaps in an exposure history arise because it may not be possible to access and measure all residences, particularly for a mobile population; previous homes may be out of the sampling area or otherwise not within the measurement protocol specifications, nonexistent, or unoccupied. Gaps in the exposure history add error to the exposure estimate.

Missing data may also result if the subject or current resident of an index house refuses to permit Rn testing in the home. The investigator must therefore appropriately increase the number of subjects to retain the desired power.

For a case-control study of lung cancer, surrogate respondents will often be required. These people may be less able to provide accurate residential history information, thereby increasing the error in the estimate of Rn exposure.

[†] Personal communication (1989), Dr. J. D. Boice, Chief, Radiation Epidemiology Branch, National Cancer Institute, Rockville, MD 20892.

SAMPLE SIZES FOR STUDIES OF Rn EXPOSURE

In designing an epidemiologic study to describe the effect of home exposure to Rn progeny, methodologic problems must be anticipated that may seriously depress the study's statistical power (Table 2). In this section, we specifically consider the effect of choice of the exposure response relation that is to be detected, of failing to account for mobility patterns in the estimation of cumulative Rn exposure, and of misspecifying the distribution of Rn concentrations in homes. In a subsequent section, we consider the effects of random error in estimates of Rn exposure.

For simplicity, we consider a hypothetical case-control study that includes all eligible male lung cancer cases between the ages of 65–69 y and twice as many controls. The study will be designed to have 90% power to reject a null hypothesis (based on a one-sided 0.05-level test) if the alternative is true. To compute power for a fixed sample size, we assume the regression model for disease outcome given by expression 1 or expression 2 and use sample size methods which accommodate continuous exposures (Lubin and Gail 1990).

Application of the sample size formulae in Lubin and Gail (1990) requires information on the distribution of exposure in the population. Nero et al. (1986) found that Rn exposure concentrations in the living area of 552 single-family dwellings from across the United States were approximately lognormally distributed, with a geometric mean of 33.3 Bq m^{-3} and a geometric standard deviation of 2.84. Assuming people reside in their homes 70% of the time and an equilibrium fraction of 50%, the geometric mean corresponds to an exposure of approximately 0.18 WLM y^{-1}. Because total cumulative exposure for 65 y olds (excluding the most recent 5 y to account for latency) would generally be estimated as 60 times the annual exposure, the distribution of cumulative exposure is approximately lognormal, with geometric mean 10.8 WLM. (For the lognormal distribution, the geometric standard deviation remains 2.84. The arithmetic mean is about 19 WLM.) For ease of calculation, cumulative exposure has been categorized into whole units and approximate formulae used.

Table 3 shows sample sizes needed to test the three hypotheses described in Table 1. For example, 251 cases and 502 controls would be needed to reject a null hypothesis of no increased risk of lung cancer, assuming β = 1.5% per WLM. (In a study with equal numbers of cases and controls, a sample of 343 cases and 343 controls is required.) If the goal were to reject β = 1.5% when the true effect is 0.75% (hypothesis 2), then 1610 cases and 3221 controls must be enrolled.

A study designed to reject an additive relative risk model for Rn and smoking, if the true model is multiplicative, requires m = 764 cases and n = 1527 controls. Ancillary calculations not shown in Table 3 indicate that m = 1220 cases, and twice the number of controls, are required to reject an additive model with a submultiplicative alternative with θ = 0.8 and 3247 cases to reject with a submultiplicative alternative with θ = 0.5.

Table 3. Effect of magnitude of risk and mobility pattern on number of required cases and on power for several tests of hypotheses (Table 1). Study based on a control-to-case ratio of 2.

		Test of hypothesis (Table 1)					
		(1)		(2)		(3)	
β	Mobility pattern	Cases	Power*	Cases	Power	Cases	Power
0.015	60 y	251	0.90	1,610	0.90	764	0.90
	3 × 20 y	529	0.66	3,156	0.67	1,906	0.65
	6 × 10 y	938	0.46	5,399	0.48	4,124	0.50
	10 × 6 y	1446	0.35	8,165	0.37	7,708	0.44
0.010	60 y	424	0.90	2,512	0.90	1,292	0.90
	3 × 20 y	952	0.63	5,163	0.66	3,344	0.64
	6 × 10 y	1736	0.43	9,060	0.46	7,326	0.49
	10 × 6 y	2723	0.32	13,907	0.35	13,776	0.43
0.005	60 y	1199	0.90	6,295	0.90	3,581	0.90
	3 × 20 y	2943	0.60	14,814	0.62	9,947	0.61
	6 × 10 y	5583	0.39	25,988	0.42	22,348	0.46
	10 × 6 y	8948	0.29	40,865	0.31	42,573	0.40

* Power for the sample size using exposure based on a single residence for 60 y. Exposure concentration distribution is given by Nero et al. (1984). For example, if in a study with 251 cases and 502 controls all subjects lived 10 y in each of six different residences, then the power to reject hypothesis 1 is 0.46.

Computations of sample sizes in Table 3 for hypotheses 1 and 2 ignore the effects of cigarette smoking. One approach for taking smoking into account tests β under the joint model of eqn (2). If we assume the joint model is multiplicative (θ = 1 in model 2) and smoking and Rn exposure are independent, then 335 and 2101 cases and twice as many controls are required for testing hypotheses 1 and 2, respectively, in contrast to the 251 and 1610 cases found previously. If Rn exposure and smoking are jointly submultiplicative, then greater numbers of subjects are required; for example, if θ = 0.8 in model 2, then 421 and 2612 cases and twice as many controls are required for testing hypotheses 1 and 2.

FACTORS INFLUENCING COMPUTED SAMPLE SIZES

The sample sizes calculated in the previous section were based on assumptions, which are unlikely to be true in most practical situations and lead to over-optimistic sample sizes. In this section we consider several features that can influence sample size computations.

Uncertainty in trend

The value of the excess relative risk parameter, β = 0.015, was taken from a BEIR IV analysis that used external population rates (NAS 1988). However, this β is itself an estimate and subject to uncertainty. In an analysis using an internal standard, the estimated exposure effect was β = 0.0134. The "time-since-exposure model" that was ultimately adopted by the BEIR IV Committee

was more complex than our model 1. For β in model 1, one might take the age-specific exposure effect from the final BEIR IV model, that is, $\beta = 0.010$. Other investigators, based on summaries of results from published papers, have developed other estimates, ranging from 0.010 to 0.0228 (ICRP 1987; Lubin 1988; Thomas et al. 1985).

The effect on the required sample size of reducing the alternative of interest is shown in Table 3. If the relative risk trend is taken to be $\beta = 0.010$ rather than 0.015, then 424 cases (and 848 controls) are needed to reject hypothesis 1, 2512 cases to reject hypothesis 2 (under the alternative, $\beta = 0.005$), and 1292 cases to reject hypothesis 3. Even greater numbers are needed if the trend parameter is $\beta = 0.005$.

Sample sizes with changes of residence

For our illustration, cumulative lifetime Rn exposure distribution was estimated as 60 times the current yearly rate of home exposure, where the rate distribution was based on the lognormal distribution estimated by Nero et al. (1986). This calculation erroneously assumes that people spend a lifetime in a single home. In fact, U.S. residents move approximately once every 5 y (U.S. Bureau of the Census 1989). Because the U.S. population is so mobile, cumulative exposures to Rn tend toward the mean, and variability among individuals decreases. Consequently, larger sample sizes are needed to detect effects.

To illustrate, suppose X is the Rn exposure rate in WLM y^{-1} and that $\log(X)$ is normally distributed with parameters $\mu = \log 0.18 = -1.71$ and $\sigma^2 = (\log 2.84)^2 = 1.09$. The distribution of subjects who live 60 y at the same residence has mean 18.6 WLM and variance $\mathrm{var}(60X) = 60^2\mathrm{var}(X) = 690.7$. Consider three alternative residency patterns: a subject living 20 y at each of three residences, a subject living 10 y at each of six residences, and a subject living 6 y at each of 10 residences. We make the simplifying assumption that concentrations of Rn in different homes occupied are statistically independent. Suppose a subject lives 20 y at each of three residences with exposure rates X_1, X_2, and X_3, where X_1, X_2, and X_3 are independent and identically lognormally distributed. Then cumulative exposure is $20X_1 + 20X_2 + 20X_3$, with mean 18.6 WLM and variance $\mathrm{var}(20X_1 + 20X_2 + 20X_3) = 230.2$. For a subject living 10 y at each of six residences, with rates $X_1, \ldots, X_6$, the mean exposure is 18.6 WLM and the variance is 115.1. Similarly, for a subject living 6 y at each of 10 residences, the mean is again 18.6 WLM, while the variance is 69.1.

The distribution of cumulative Rn exposure with multiple residences is the sum of lognormal distributions but is not itself lognormal. The distributions for the different mobility patterns that were described above are shown in Fig. 1. The modes of the distributions increase toward the mean as the number of residences increase, while the skewness and variance decrease.

The influence of mobility on sample size and attained power is shown in Table 3 for each of the hypotheses of interest and for various β values. Mobility increases the required number of cases and controls from twofold to as much as 10-fold. If mobility is ignored, then the true power can be substantially less than the assumed 90% power.

Misspecification of Rn *concentration distribution*

The estimate of the distribution of Rn rates in homes by Nero et al. (1986) was based on a limited number of surveys from diverse sources and thus may not represent the true distribution in the U.S. or in a specific study area. Indeed, the authors acknowledged that Midwestern and Southern areas of the country may be underrepresented.

Two other national data sources for Rn rates have been reported. In a survey of university physics professors, Cohen found that the data were consistent with a lognormal distribution. He estimated a geometric mean of 38.1 Bq m^{-3} and a geometric standard deviation of 2.36 (Cohen 1986), very similar to the estimates (33.3 Bq m^{-3} and 2.84) reported by Nero et al. (1986).

A third set of data was reported by a commercial enterprise, which collected over 61,000 measurements from about 30,000 homes (Alter and Oswald 1987). The geometric mean was 69.9 Bq m^{-3} and the geometric standard deviation was 4.30. It is difficult, however, to assess the representativeness of this data source, since measurements derive from concerned homeowners who actively sought information about Rn levels, since measurements may not have been from living areas, and since the season or seasons of measurement are not reported.

It is important to note that the sources of data on exposure concentrations are based on national surveys, whereas case-control studies are carried out within limited geographic areas. If the case and control ascertainment region is narrowly drawn, then the range of Rn concentrations and hence cumulative Rn exposure among cases and controls may be smaller than national data indicate, requiring an increased sample size for detecting excess risk. On the other hand, case-control studies may be carried out in areas where there is a potential "Rn problem" and thus where variation could be greater than nationally.

Table 4 provides some indication of the effects of various exposure concentration distributions on study power. The required sample size increases considerably as the variance decreases. For hypothesis 1, 251 cases are required if the geometric mean and geometric standard deviation are 10.8 and 2.84, respectively, in comparison to 497 cases if the geometric mean is the same and the geometric standard deviation is 2.30. Table 4 also shows that an increase in variance of the exposure distribution (by increasing the geometric mean or the geometric standard deviation) reduces required size. Until representative surveys are completed, it is difficult to know how accurately the national estimates characterize concentrations throughout the country or within specific regions.

Testing hypothesis 1 *when* Rn *exposure is measured with error*

In the preceding calculations, exposure estimates were assumed to reflect true levels of exposure without

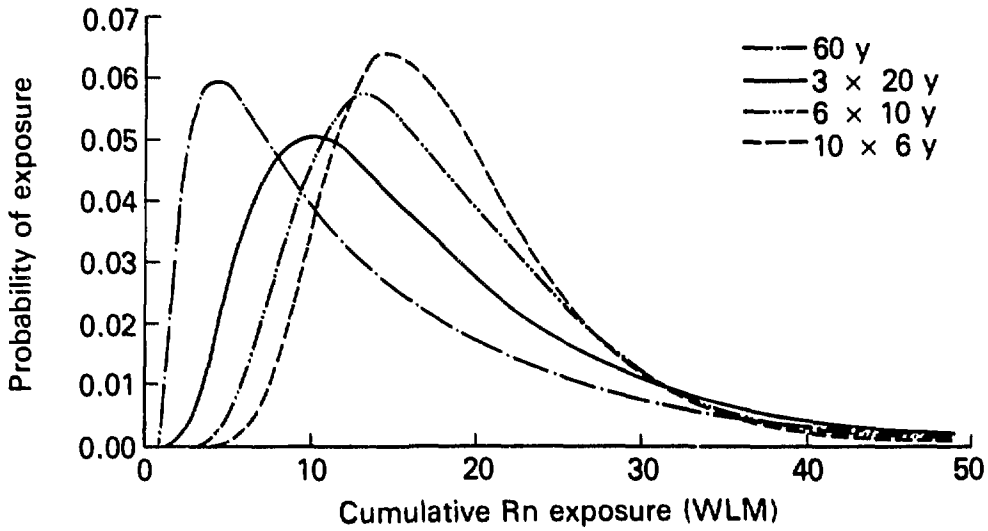

Fig. 1. Distribution of 60 y of cumulative exposure to Rn and its decay products under various residency patterns.

error. We now consider the impact of imprecise measurement of exposure.

The Appendix develops alternative formulae for determining sample size for hypothesis 1, when exposure W is subject to error. We assume that true cumulative exposure W in WLM is lognormal and that $\log(X)$ has parameters $\mu = -1.71$ and $\sigma^2 = 1.10$. However, W cannot be observed directly. Instead, we observed $Y = W \times U$, where U is a multiplicative error. The error U is assumed to be lognormally distributed, with $\log(U)$ having parameters 0.0 and $\sigma^2 f^2$. The factor f determines the error as a proportion of the standard deviation of the true exposure, as measured on the log scale. The error U has median one and mean $\exp(0.55 f^2)$.

Table 4. Effect on sample size and power of altering the distribution of rates of Rn decay products in the home. Results based on subjects living 60 y at a single residence and on $\beta = 0.015$. Geometric mean in units of cumulative WLM.

Cumulative Rn distribution		Test of hypothesis					
		(1)		(2)		(3)	
GM[a]	GSD	Cases	Power[b]	Cases	Power[b]	Cases	Power[b]
6.0	2.84	568	0.66	3087	0.68	1625	0.66
8.4	2.84	364	0.81	2071	0.82	1029	0.82
10.8	2.84	251	0.90	1610	0.90	764	0.90
13.2	2.84	213	0.94	1316	0.94	591	0.95
15.6	2.84	177	0.97	1143	0.96	491	0.97
10.8	2.30	497	0.70	2883	0.71	1622	0.68
10.8	3.46	188	0.96	1091	0.97	458	0.98

[a] Using the conversion 37 Bq m^{-3} = 0.2 WLM y^{-1}, geometric mean values correspond to Rn concentrations of 18.5, 25.9, 33.3, 40.7, and 48.1 Bq m^{-3}, respectively.

[b] Power relative to a study with 251, 1610, or 764 cases and twice the number of controls, respectively.

Table 5 shows that the number of cases and controls that are needed to reject hypothesis 1 increases as the degree of error, f, increases. With exposure misspecified by 50%, about 1.5 times the number of cases and controls are needed to retain the specified power to reject the null. Even greater numbers are needed when both exposure errors and mobility patterns are considered.

We can only speculate about values for f in a real study situation. However, given the multitude of factors that may contribute to error in estimating exposure for an individual, it is likely that errors exceed 50%. More research in the area of exposure estimation is needed.

(When exposure is measured with error, the observed exposure response relationship is no longer linear but concave from below, the product of the true β and a decreasing function of the observed exposure Y. A plot of the probability of disease given the observed Y is shown in the Appendix. With increasing error rate, the relative risk pattern becomes increasingly curvilinear.)

DISCUSSION

Many case-control studies of domestic Rn exposure and lung cancer are now in progress throughout the world. These studies are being conducted to test whether Rn causes lung cancer in the general population and to determine the associated levels and modifiers of risk. The results of these studies will supplement and may possibly replace the data from underground miners as the basis for choosing acceptable levels of indoor exposure. The evidence from these studies must be viewed within the context of the uncertainties and power limitations which we have discussed, so that the public can be realistically advised about Rn. For sound public policy to evolve, the uncertainty concerning the quantitative risks of Rn exposure in homes must be minimized.

Table 5. Effect of measurement error, f, and mobility pattern on sample size required to reject no trend with Rn exposure, $\beta_0 = 0$, when the true trend is $\beta_1 = 0.015$. Study based on a control-to-case ratio of 2.

	Mobility pattern							
	60 y		3 × 20 y		6 × 10 y		10 × 6 y	
f	Cases	Power*	Cases	Power*	Cases	Power*	Cases	Power*
0.00	251	0.90	529	0.66	938	0.46	1,446	0.35
0.30	288	0.86	656	0.58	1,303	0.37	2,303	0.26
0.50	365	0.79	916	0.48	2,050	0.28	4,059	0.19
1.00	973	0.48	2,987	0.23	8,002	0.14	18,032	0.10
1.50	4,186	0.22	13,934	0.12	39,456	0.09	91,875	0.07
2.00	29,542	0.12	100,308	0.08	287,644	0.07	674,540	0.06

* Power relative to a study with 251 cases, 502 controls, and no error in exposure, $f = 0.0$.

Analyses we presented show that mobility, error in estimation of Rn exposure, and other factors can influence sample size computations. However, our calculations depended on certain modeling and exposure distribution assumptions. Because some of our assumptions may not apply generally, computed sample size for any particular study may differ from those shown in the tables, although the impact of the issues we discuss is still relevant.

Our analyses suggest that case-control studies of indoor Rn and lung cancer may require substantial numbers of subjects, and that realistically such studies may never be able to answer many of the subtle questions about risk patterns that burden current risk assessments with uncertainty. Even the most carefully designed and conducted investigations are subject to substantial error in dosimetry, particularly when used to estimate temporally remote exposures. Inappropriate design assumptions with regard to the underlying effect, subject mobility, and exposure distribution also seem inevitable.

If the effect of domestic Rn exposure (per unit WLM) is greater than current estimates, then the required sample sizes will be smaller than those in Tables 3 and 5. There are two reasons to believe that the exposure-response relationship in homes might be steeper than that derived from miner data. First, dosimetry was uncertain for miners (NAS 1988). If Rn exposures for the miners were estimated with substantial error, then the estimated exposure-response relationship may be too low, so that the true β may be higher than the 0.015 that we have assumed. Second, our calculations were based on cumulative exposure only and did not allow for effects of exposure rate. Although the effects of exposure rate on lung cancer risk are still unclear (NAS 1988), two studies have suggested that for fixed cumulative exposure, prolonged duration at low concentrations may be more deleterious than short duration at high concentrations (Lubin et al. 1990; Sevc et al. 1988). Since Rn concentrations are generally lower in homes than in mines, domestic Rn exposure may have greater effect per WLM than in mines.

One other source of uncertainty relates to the fact that application of miner-based risk models to home exposures requires the conversion of indoor measurements of Rn (Bq m^{-3}) to levels of Rn progeny (WLM). For homes, an estimate of 50% of equilibrium for Rn with its progeny is usually assumed. Recent work, however, suggests that the equilibrium fraction for homes may be much lower, perhaps 20% (Toohey et al. 1987), and quite variable. Thus, domestic exposure (in WLM) for our sample size and power calculations may need to be reduced by a factor of 0.4 (=0.20/0.50). Alternatively, one can modify β, so that $\beta = 0.015 \times 0.40 = 0.006$ replaces 0.015. For example, if the equilibrium factor is 20%, as opposed to 50%, 899 cases and 1,798 controls are required for testing hypothesis 1 (assuming 60 y residence in a single home), in contrast to 251 cases and 502 controls.

Only very large case-control studies can substantially reduce the uncertainty of extrapolating from mines to homes (hypothesis 2). Calculations suggest that large numbers of subjects are required to reject miner-derived risk models. More extreme alternative hypotheses, for example, a null β of 1.5% and an alternative of 0.15% per WLM, reduce sample size requirements but seem unlikely. Current dosimetric analyses have provided a covergent view that any differences in the potency of Rn as a carcinogen in the two environments are likely to be slight (ICRP 1987; NAS 1988; NCRP 1984).

In our illustrations, lifetime exposure was used to calculate sample size. For feasibility, some studies measure Rn concentrations and estimate exposures only for the recent past. Analyses of miners provide some justification for this approach, since risks decline with time since exposure (NAS 1988). However, if earlier exposures do in fact contribute to risk, omitting them by design in effect builds in noise to the exposure estimate (see Table 5). If, on the other hand, earlier exposures are biologically irrelevant, then cumulating over a shorter interval reduces both the exposure and its variability and consequently increases required sample size. For example, testing hypothesis 1 using 30 y of exposure (the period from 5 y to 35 y prior to interview) requires 678 cases and 1,355 controls (assuming continuous residence in a single home), in contrast to the 251 cases and 502 controls calculated previously (Table 3).

Our analyses focused on problems arising from inappropriate assumptions in designing a study and from errors in estimating exposure to Rn or to Rn decay prod-

ucts. However, other methodological constraints may also affect a case-control study of Rn and lung cancer. For example, we did not consider the uncertainty introduced by errors in the ascertainment of tobacco consumption (Table 5). Misclassification of cases may also lower study power, while differential misclassification may distort results (Garfinkel et al. 1985; Gordis 1982; Lerchen and Samet 1986; McFarlane et al. 1986, 1987; Pickle et al. 1983).

We thus urge cautious interpretation and reduced expectations for case-control studies. Predictably, some of the studies now in progress will not show increased lung cancer risk in association with estimated exposure to Rn; the negative findings of these studies may only reflect low study power from insufficient sample size, as well as multiple sources of bias acting to reduce effects towards the null. In these instances, we urge investigators to recompute study power at the completion of a study, based on the effective sample size realized, the observed exposure distribution in controls, and, if possible, estimates of measurement error. Some of these studies or some subgroups of a study population may show effects of Rn exposure that are substantially above levels expected based on miner studies; these findings should also be carefully considered as they too may be the result of small sample sizes or may arise due to the selective publication of positive results (Land 1980). For studies of low power, confidence intervals will therefore be more meaningful than point estimates of effects.

Because of the design issues that we have discussed, we suggest the pooling of data from studies with the largest numbers of cases and with the most precise estimates of Rn exposure as the best, and perhaps the only, approach for addressing present research needs. We also advise that careful consideration be given these difficult design problems before initiating further case-control studies of lung cancer and Rn exposure.

Acknowledgments—The authors wish to thank Dr. John D. Boice, Jr. for useful discussions on these and related topics and for editorial comments. Dr. Samet was supported in part by U.S. Department of Energy, Office of Energy Research, under grant No. DE-FG04-86ER60452.

REFERENCES

Alter, H. W.; Oswald, R. A. Nationwide distribution of indoor radon measurement: A preliminary data base. J. Air Pollut. Control Assoc. 37:227–231; 1987.

Bierma, T. J.; Croke, K. G.; Swartzman, D. Accuracy and precision of home radon monitoring and the effectiveness of EPA monitoring guidelines. J. Air Pollut. Control Assoc. 39: 953–959; 1989.

Borak, T. D.; Johnson, J. A. Estimating the risk from inhalation of radon daughters indoors: Review and evaluation. Las Vegas, NV: Environmental Protection Agency; EPA 600/6-88/008; 1988.

Cohen, B. L. A national survey of ^{222}Rn in U.S. homes and correlating factors. Health Phys. 51:175–183; 1986.

Cohen, B.; Gromicko, N. Variation of radon levels in U.S. homes with various factors. J. Air Pollut. Control Assoc. 38:129–134; 1988.

Crameri, R.; Brunner, H. H.; Buchli, B.; Wernli, C.; Burkart, W. Indoor Rn levels in different geological areas in Switzerland. Health Phys. 57:29–38; 1989.

Fleischer, R. L.; Mogro-Campero, A.; Turner, L. G. Indoor radon levels in the Northeastern U.S.: Effects of energy-efficiency in homes. Health Phys. 45:407–412; 1983.

Gail, M. H.; Wacholder, S.; Lubin, J. H. Indirect correction for confounding under multiplicative and additive risk models. Am. J. Ind. Med. 13:119–130; 1988.

Garfinkel, L.; Auerbach, O.; Joubert, K. Involuntary smoking and lung cancer: A case-control study. J. Natl. Cancer Inst. 75:463–469; 1985.

George, A. C.; Breslin, A. J. The distribution of ambient radon and radon daughters in residential buildings in the New Jersey-New York area. In: Gesell, T. F.; Lowder, M. W., eds. Springfield, VA: National Technical Information Service, Natural Radiation Environment III/DOE CONF-780422; 1980:1271–1292.

Gordis, L. Should dead cases be matched to dead controls? Am. J. Epidemiol. 115:1–5; 1982.

Hess, C. T.; Fleischer, R. L.; Turner, L. G. Field and laboratory tests of etched track detectors for ^{222}Rn: Summer-vs-winter variations and tightness effects in Maine houses. Health Phys. 49:65–79; 1985.

Hess, C. T.; Weiffenbach, C. V.; Norton, S. A. Environmental radon and cancer correlations in Maine. Health Phys. 45: 339–348; 1983.

International Commission on Radiation Protection. Lung cancer risk from indoor exposure to radon daughters. A report of a task group. Oxford: Pergamon Press; ICRP Publication No. 50; 1987.

James, A. C. Lung dosimetry. In: Nazaroff, W. W.; Nero, A. V., eds. Radon and its decay products in indoor air. New York: J. Wiley & Sons; 1988:259–309.

Jönsson, G. Indoor ^{222}Rn measurements in Sweden with the solid-state nuclear track detector technique. Health Phys. 54: 271–281; 1988.

Land, C. E. Estimating cancer risks from low doses of ionizing radiation. Science 209:1197–1203; 1980.

Lerchen, M. L.; Samet, J. M. An assessment of the validity of questionaire responses provided by a surviving spouse. Am. J. Epidemiol. 123:481–489; 1986.

Lubin, J. H. Models for the analysis of radon-exposed populations. Yale J. Biol. Med. 61:195–214; 1988.

Lubin, J. H.; Boice, J. D., Jr. Estimating radon-induced lung cancer in the U.S. Health Phys. 57:417–427; 1989.

Lubin, J. H.; Gaffey, W. Relative risk models for assessing the joint effects of multiple factors. Am. J. Ind. Med. 13:149–167; 1987.

Lubin, J. H.; Gail, M. H. On power and sample size for studying features of the relative odds of disease. Am. J. Epidemiol. 131:552–566; 1990.

Lubin, J. H.; Gail, M. H.; Ershow, A. G. Sample size and power for case-control studies when exposures are continuous. Stat. Med. 7:363–376; 1988.

Lubin, J. H.; Qiao, Y. L.; Taylor, P. R.; Yao, S. X.; Schatzkin, A.; Xuan, X. Z.; Mao, B. L.; Rao, J. Y.; Li, J. Y. A quantitative evaluation of the radon and lung cancer association in a case-control study of Chinese tin miners. Cancer Res. 50: 174–180; 1990.