Radon and lung cancer: case-control *vs.* ecologic studies (*)

H. FRIEDMANN
Institut für Radiumforschung und Kernphysik, University of Vienna
Boltzmanngasse 3, A-1090 Vienna, Austria

(ricevuto il 9 Giugno 1998; revisionato il 9 Dicembre 1998; approvato il 18 Dicembre 1998)

Summary. — Problems in epidemiological studies in connection with the determination of radon-caused lung cancer are discussed. Biasing and confounding effects have led to very divergent results in epidemiological studies and there is still doubt on the carcinogenic effect of radon progenies in the concentration range of usual indoor air (say below several hundred Bq/m^3). Nevertheless, epidemiological investigations seem to be the key tool to decide this question. Main advantages and disadvantages of case-control and ecologic studies are discussed and an overall uncertainty estimation is given for both types of investigations.

PACS 87.50 – Effects of radiation and external fields on biomolecules, cells and higher organisms.
PACS 87.50.Gi – Ionizing radiations (ultraviolet, X-rays, ions, γ-rays, electrons, positrons, neutrons, and mesons, etc.).
PACS 89.60 – Environmental studies.
PACS 87.90 – Other topics in biological and medical physics (restricted to new topics).
PACS 01.30.Cc – Conference proceedings.

1. – Introduction

Epidemiological investigations on miners (cohort studies) clearly showed an enhanced lung cancer risk for people working in an atmosphere with high radon progeny concentration (a summary can be found in [1] and [2]). From these investigations a risk coefficient was deduced and according to the usual philosophy in radiation protection the observed risk in the miner studies was linearly extrapolated to the radon (progeny) concentration people are commonly exposed to in their homes. Such an extrapolation leads to a radon-related contribution of 5 to 15% (depending on the mean radon concentration in a country) to the total lung cancer mortality. With a normalised total lung cancer death rate of approximately 50 per 100000 inhabitants in a year, radon is responsible for the lung cancer death of approximately 50 people per

(*) Paper presented at the "Fourth International Conference on Rare Gas Geochemistry", Rome, October 8-10, 1997.

year in a population of 1 million. The question is whether the linear extrapolation of the lung cancer risk from the miners' studies to the average population is justified. Primarily, it is important to know if an additional lung cancer risk at lower radon concentrations, below several hundred Bq/m^3, say, exists at all. If this can be proved then the question arises whether the risk is as large as it appears from a linear extrapolation or rather lower. Supposing the latter is true, mitigation measures should start in houses with high radon exposure of the inhabitants while in case of a linear dose-effect relation a general reduction of indoor radon concentrations seems to be necessary.

In radiation protection the dose concept is generally accepted. That means that any radiation exposure (in case of radon measured in Bq·h/m^3, WLM, mJ·h/m^3, …) is converted into a dose by a dose conversion factor (DCF). The dose multiplied by a detriment factor finally gives the health risk. Within this concept the problem is to find the correct conversion factors and their possible exposure dependencies. Three ways are thinkable to solve these questions:

– The theoretical approach: The interaction of radon progenies with cells can be estimated by mathematical models and DCFs can be deduced [3]. Unfortunately there is a discrepancy between the DCF deduced from miners data and from modelling. The latter gives substantially higher DCFs.

– Animal experiments: It is rather easy to expose animals to a certain radon progeny dose and to observe the lung cancer risk. However, only with rather high radon exposures a statistically significant effect can be observed. This is, of course, also a consequence of the limited number of animals in such experiments. Another main question is whether results from animal experiments can be transferred to humans. This is not only due to anatomic differences, but also to the length of exposure, age at exposure and possible dose-rate effects (see, *e.g.*, [4]).

– Epidemiological investigations.

2. – Problems in epidemiological investigations on the lung cancer risk of radon

The essential problem in epidemiological studies on the lung cancer risk of radon is to detect a very small effect above a large background. This means that it is necessary to have a large data base to reduce statistical uncertainties and to have a very good knowledge of the background (lung cancer caused by other factors) as well as a good understanding of the reasons causing the investigated effect (lung cancer). Deficiencies in one of these groups can totally destroy any correlation. This is probably the reason for the very controversial results of epidemiological investigations on radon and lung cancer [2, 5].

Generally, there are two epidemiological approaches: the case-control study and the ecologic study. The case-control study tries to estimate the radon exposure of people with lung cancer and compares it with the exposure of people from a control group without lung cancer. The control group has to fit in most of the relevant parameters (age, gender, smoking etc.) with the investigated lung-cancer group. The ecologic study compares the lung cancer rate in certain areas as a function of the mean radon concentrations in these areas. Both types of investigations suffer from different shortcomings, some concerning both types of investigations and others specifically one type.

2·1. *Common problems.* – The main problems are the insufficient control of other lung-cancer–causing factors (confounders) and unrealistic uncertainty estimation. Both problems may have several reasons. There are a number of avoidable error sources which sometimes were not taken into consideration as seriously as necessary. These are mainly:

– standardisation to different reference groups (age distribution, gender, urbanisation, social distribution etc.);

– usage of surrogates (γ-dose rates, geological information etc.) instead of real radon data;

– incorrect measurement results and wrong annual mean estimates;

– too small database.

Apart from this, there are a lot of problems which directly influence the outcome of epidemiological studies but are not yet solved or the used assumptions are not really proved. So many parameters and/or the way they influence the lung cancer risk are unknown, *e.g.*:

– Does radon act as an absolute (independent of the background lung cancer risk) or as a relative risk (increasing with background lung cancer risk)?

– Is there a dose-rate dependence of the radon risk?

– Is there an age dependence of the radon risk?

– What is the importance of the age at exposure?

– How can the aerosol-size distribution influence the radon risk and how can aerosol-size distribution be determined for the past?

– How to determine the annual mean ratio of unattached to attached progenies for all investigated groups/persons, especially for the past?

– If radon gas is measured: what is/was the relevant equilibrium factor?

– What is/was the breathing volume of the investigated groups?

– Do we have enough information about time-dependent changes in confounding influences (*e.g.*, change in smoking behaviour)?

These problems may have similar or different influence on case-control and ecologic studies but generally concern both types of investigations. But there are a lot of problems which mainly concern only one type of epidemiological studies.

2·2. *Main problems in case-control studies.* – One group of main problems in case-control studies arise from the fact that individual data are necessary and no averaging is possible. Uncertainties are often introduced when the data from the lung cancer group have to be retrieved from relatives and friends. Further problems result from fitting or at least normalising the control group to the lung cancer group. Finally, several assumptions are necessary during the analysis of the data, to model the influence of radon and other factors on the lung cancer risk (model dependence). Essential questions are:

– What is the uncertainty in the determination of the radon (progeny) exposure for a whole life (or at least for the last 20 years)? Is it possible to determine the radon

progeny exposure in former homes, at work etc., and all the relevant parameters in a way that gives sufficient information for a risk estimate?

– How to weight the exposure in the past? According to the miners data it is assumed that the exposure from more than 15 years ago is half as important as the exposure incurred between 5 and 15 years before risk evaluation [4]. But this remains a model assumption for the common population and the uncertainty in this assumption is rather difficult to estimate.

– What personal lung cancer risks are known and to what extent may they influence the radon lung cancer risk? How did these additional risks (smoking, asbestos, dust etc.) change in time and how are former exposures to such confounding factors to be weighted?

– Is it possible to identify all special events which may have influenced the lung cancer risk for members of the investigated groups, *e.g.*, diseases?

– The control group must fit the lung cancer group in gender, age distribution, smoking behaviour, social behaviour etc. This is usually not possible. Therefore, modelling of the influence of all the relevant parameters is necessary to normalise both groups to the same parameter set. Is it possible to estimate the uncertainty introduced by this normalisation correctly?

– How can the risk of confounding factors be measured? *E.g.*, is pack-years an adequate measure for smoking? It does not take into account tar concentration, life-time variation of smoking behaviour, way of smoking (inhaling) etc.

Several of the above questions can be answered only by using the information available from an average of people. Therefore, even in case-control studies some of the severe problems of ecologic studies are present.

A severe non-scientific problem with case-control studies is their costs which are generally higher than for ecologic studies.

2·3. *Main problems in ecologic studies.* – In ecologic studies individual parameters are substituted by mean values of a group. This can reduce uncertainties for several variables, however, a lot of new problems will be introduced.

Ecologic studies can be divided into 2 classes: comparison studies and regression studies. In a comparison study the lung cancer rate in two or more areas are compared with the mean radon exposure in these areas. In a regression study the population of an area is divided into groups with different radon exposure and the total lung cancer risk is then calculated as the sum of the risks in these groups:

$$P = \alpha_0 + \alpha_1 S(R_1) + \alpha_2 S(R_2) + \alpha_3 S(R_3) + \ldots + \alpha_n S(R_n)$$

with P the total lung cancer risk, $S(R_i)$ the size of the group with a mean radon (progeny) concentration R_i and α_i the lung cancer risk related to the mean radon concentration R_i. The α_i can be computed by least-squares fit or by maximum-likelihood methods or some other methods from the data of more than n regions. Usually, the frequency distribution of the indoor radon concentration follows a log-normal distribution. Therefore, $S(R_i)$ can be calculated from a determination of the radon concentration in a statistically correctly selected sample. The problem is primarily bound to α_0 (background) which can be expected to be different in different areas and must be adjusted for all other known lung cancer risks. If synergy effects

between the radon risk and other risks exist, the α_i are dependent on α_0, which makes any regression analysis even more complicated.

There are a lot of problems in ecologic studies:

– The uncertainty in the determination of the (annual, life-time) exposure is dependent on the applied measurement method (long-term measurement, repeated measurements) but should not be as critical as in case-control studies.

– Uncertainties in the frequency distribution of the indoor radon concentration (usually log-normal distribution) can lead to wrong risk estimates. If the lung-cancer risk is not linearly dependent on the radon exposure, equal mean radon values may result in significant different lung cancer risks because of different geometric standard deviation. This problem cannot easily be solved in a comparison study.

– It must always be proved that the sample selection from an area is random.

– How are other lung cancer risks and effect modifiers (non-additive risk effects) to be treated? Mean values may not be representative for the risk ('non-linearity' — *e.g.*, many people smoking only few cigarettes and few people smoking many cigarettes give similar means). Profound investigation on the social distribution of confounding effects (*e.g.*, in Austria smoking is strongly correlated with urbanisation and several other parameters) and its change in time (increase in women lung cancer rate because of increased smoking) is necessary.

– Wrong or insufficient cancer statistics: Tests of reliability of the cancer data are necessary. Large errors can be introduced in any statistics if the lung cancer statistics are incorrect even in a few areas only. Special attention must be paid to persons who died in hospital counting them to their domicile.

– Migration: There is always a migration between areas of different mean radon concentration. In principle it should be possible to correct for this effect on the basis of census data and data of the registration office. (In Europe the migration effect is of the order of equal or less than 5% per year and is not as important as in the U.S.A.) In addition, the time dependence of the radon risk (age at risk) could play an important role in case of migration.

– A model is necessary to correct for confounding effects. The errors introduced by wrong modelling of the lung cancer risk of other noxes depend on the size of their risk. Synergy effects can lead to substantial errors because the mean of the product of two effects is different from the product of the means ($\langle XY\rangle \neq \langle X\rangle\langle Y\rangle$). This problem can partly be overcome by a careful selection of areas with very similar influence of confounding factors.

– A correct estimate of the uncertainties (which is sometimes rather difficult) and sufficient statistical power are always necessary. This problem is of course similar for case-control studies.

3. – Discussion

Today case-control studies are favoured because it is assumed that biasing can be more easily controlled. Several case-control studies are already launched [6-10] and pooling of these data is planned for the future [11]. There are publications, *e.g.* [12], which tried to show the on-principle inability of ecologic studies to estimate the radon-caused lung cancer risk. Most of these papers do not take uncertainties into

account. However, the essential point in any epidemiological study is the correct estimate of uncertainties in the investigation. This means that the size of a possible radon-caused lung cancer risk which can be detected within an epidemiological study can only be estimated by a strict analysis of the uncertainties of all possible input parameters, an analysis of the influence of all necessary assumptions and normalisation procedures, and a correct statistical treatment of all data.

For a typical case-control study with 2 groups of approximately 1000 people each, perfect cancer diagnosis and a very well-matched control group, the total uncertainty (1 standard deviation) for the relative change in lung cancer risk can be estimated to be of the order of 70% (contributions ordered by importance: model-dependent correction for other risks, uncertainty in exposure, synergy effects, matching). This means that a really significant result can only be expected for groups when the ratio of their radon concentrations exceeds a factor of two.

For a comparison study with 10000 inhabitants in each region, a 10 years' perfect cancer statistics and a lung cancer rate of the order of 5 per year for the investigated areas, the total uncertainty is again in the 70% range (contributions ordered by importance: model-dependent correction for other risks, migration, synergy effects). In this case a significant result can only be expected in regions with radon concentrations leading to an odds ratio of more than 2.

There are some advantages in case-control studies, however, a correct ecologic study will give comparable results. Therefore, statements denying the importance of ecologic studies seem to be rather questionable.

REFERENCES

[1] Lubin J. H., Boice J. D., Edling C. H., Hornung R., Howe G., Kunz E., Kusiak A., Morrison H. I., Radford E. P., Samet J. M., Tirmarche M., Woodward A., Xiang Y. S. and Pierce D. A., *US National Institutes of Health*, *NIH publication*, **94** (1994) 3644.
[2] NRC, National Research Council, *BEIR VI: The Health Effects of Exposure to Indoor Radon* (National Academic Press, Washington, D.C.) 1998.
[3] ICRP, *Annals of the ICRP*, **24**, no. 14 (1994), ICRP Publ. 66.
[4] National Academy of Science, *BEIR IV: Radon and Other Internally Deposited Alpha Emitters* (National Academic Press, Washington, D.C.) 1988.
[5] Lubin J. H. and Boice J. D., *J. Nat. Cancer Inst.*, **89** (1997) 49-57.
[6] Auvinen A., Mäkeläinen I., Hakama M., Casrén O., Pukkala E., Reisbaka H., Rytömaa T., *J. Nat. Cancer Inst.*, **86** (1994) 1829-1837; **88** (1996) 966-972.
[7] Létourneau E. G., Krewski D., Choi N. W., Goddard M. J., McGregor R. G., Zielinski J. M. and Du J., *Am. J. Epidemiol.*, **140** (1994) 310-322.
[8] Pershagen G., Åkerblom G., Axelson O., Clavensjö B., Damber L., Desai G., Enflo A., Lagarde F., Mellander H., Svartengren M. and Swedjemark G. A.: *New England J. Med.*, **330** (1994) 159-164.
[9] Ruosteenoja E., Mäkeläinen I., Rytömaa T., Hakulinen T. and Hakama M., *Health Physics*, **71** (1996) 185-189.
[10] Alavanja M. C. R., Brownson R. C., Lubin J. H., Berger E., Chang J., Boice J. D. jr., *J. Nat. Cancer Inst.*, **86** (1994) 1829-1837.
[11] Samet J. M., *J. Nat. Cancer Inst.*, **89** (1997) 4-6.
[12] Greenland S. and Morgenstern H., *Int. J. Epidemiol.*, **18** (1989) 269-274.