How It *Really* Happened

Smoking and Lung Cancer

SIR RICHARD DOLL

ICRF/MRC/BHF Clinical Trial Service Unit and Epidemiological Studies Unit, Radcliffe Infirmary, Oxford, United Kingdom

The research that led Professor Bradford Hill and me to conclude that "cigarette smoking is a factor, and an important factor, in the production of carcinoma of the lung" (Doll, R., B. Hill, Smoking and carcinoma of the lung: a preliminary report, *British Medical Journal*, 1950;2:739–748) had been designed to find an explanation for the extraordinary increase in the mortality attributed to the disease over the previous 30 years. I have described elsewhere the background to the research (*Statistical Methods in Medical Research* 1998;7:87–117) and describe here only our attitude to that work and why we followed it with a cohort study of British doctors, which led to the demonstration of so many other harmful effects of smoking.

Our initial work was planned to be a case-control study of patients suspected of having cancer of the lung, stomach, or large bowel who had been admitted to 20 large London hospitals. Patients with stomach or large bowel cancer were included to enable us to distinguish between findings that might relate to cancer of the lung and those (if there were any) that related to cancer in general. Patients for whom the diagnosis was only suspected were included in order to give us enough time to learn about their admission and arrange for them to be interviewed before they were discharged, which, in those days, would probably have been two or three weeks. Diagnoses were in fact often changed; subsequently, I had to visit each hospital to review the patients' records and determine the discharge diagnoses and the strength of the evidence on which they had been based, noting for the lung cancer patients the presenting symptoms, the site of the lesion, whether the diagnosis had been confirmed histologically and, if it had been, the histological type.

Case-control studies, which have subsequently become one of the central planks of epidemiology, were not then common practice and there was no standard way of carrying them out. Our first important decision had, therefore, to be the choice of controls, and this was described in our first publication. The second was the design of the questionnaire. It was laudably short, filling only three and a half pages, but even so it included all the questions that we could think of that might be relevant to any of the three selected types of cancer. To guide us in its design we had only the knowledge that lung cancer had become progressively more common since the end of the First World War, was much more common in men than in women, and was somewhat more common in large towns than in the countryside, and that cancer could be caused in humans by ionizing radiation and the tar produced by the combustion of coal.

Correspondence and requests for reprints should be addressed to Sir Richard Doll, ICRF/MRC/BHF Clinical Trial Service Unit and Epidemiological Studies Unit, Harkness Bldg., Radcliffe Infirmary, Oxford OX2 GHE, UK.

Am J Respir Crit Care Med Vol 162. pp 4–6, 2000
Internet address: www.atsjournals.org

The most popular idea in the minds of those few people who thought about the problem at all, and who were prepared to consider that cancers might be capable of prevention instead of being an inevitable accompaniment of ageing, was that the disease was caused by the heavy pollution of the air in large towns by the combustion products of coal, which in many houses was burnt in open fires as the principal method of heating. This idea never appealed to me greatly, despite the intense smogs that sometimes caused cars to be abandoned in the street or to be led home by a pilot with a light to identify the edge of the road, for the simple reason that the amount of coal burnt each year had, if anything, slightly decreased over the previous 50 years. Similarly, the idea that tobacco was the cause was not very attractive, because the increase in the average consumption per person had been quite small, there being, it seemed, no obvious reason why the smoke from cigarettes, which had largely replaced cigars and pipes, should be so much more hazardous than the smoke from tobacco burnt in other ways. Both Bradford Hill and I smoked at the time, he almost only smoked pipes whereas I smoked moderately both pipes and cigarettes. What seemed to me the most likely cause was something to do with cars, either their exhaust fumes or perhaps particles from the new tarred surface of roads that had accompanied the development of motor transport. The only way that I could see to enable this to be tested was to record the past and present occupation of patients with and without the disease; however, this failed to suggest any particular exposure to motor traffic. As for smoking, we decided that we should have to record this in detail and as it turned out, most importantly, define a non-smoker strictly as someone who had never smoked as much as one cigarette a day for as long as one year.

Computers were not then available and I kept a list of all the patients interviewed in notebooks, ensuring that those subsequently shown to have lung cancer had suitable matched controls, while the information obtained from the interviews was coded and entered commercially on punch cards for subsequent analysis by our own Hollerith card sorter. Smoking habits were easy to analyse, and I kept a written tabulation for my own interest. It was not long before I was struck by the fact that when the patient with suspected lung cancer had been recorded as a nonsmoker, the diagnosis was seldom confirmed, the discharge diagnosis being, for example, bronchiectasis or, if malignant, fibrosarcoma of the chest wall or mesothelioma.

By the time we had data on several hundred patients it was obvious that the principal difference between the patients with and without lung cancer was their smoking habits, and we had to make up our minds whether the association was due to chance, bias, confounding, or to cause and effect. The evidence that led us to conclude that it was due to the last (and which led me to give up smoking in 1949) is described in our first paper and the lines of thought we followed were subsequently

written up by Hill and have become widely known as "Hill's guidelines." One bit of the evidence is, however, worth recording here, for we obtained it serendipitously. We had of course considered the possibility that the interviewers might be biased in recording the smoking histories, delving perhaps more energetically into the past if the diagnosis was known to be cancer of the lung, and we tried to arrange that they did not know whether the suspect diagnosis was cancer of the lung, large bowel, or stomach. The interviewers were therefore asked to record on each cancer patient's questionnaire whether the type of cancer from which the patient suffered (lung, stomach, or large bowel) was known to them when the interview was conducted. In practice it was nearly always recorded as "known," and there were too few cancer patients successfully interviewed blind to knowledge of the type of the disease, for us to be able to analyse the "blind" results separately. What we had not planned for was the large number of patients for whom the admission diagnosis of lung cancer turned out to be wrong and which had to be changed when the diagnoses were reviewed. We were consequently able to eliminate bias by showing that when the suspect diagnosis of lung cancer was disproved, the patient's smoking habits were similar to those of the controls and sharply different only when the diagnosis of bronchial carcinoma was confirmed.

Towards the end of 1949 we thought we had enough data and took our findings to Sir Harold Himsworth, who had become the Secretary of the Medical Research Council, which had funded our research. Sir Harold was a distinguished clinical scientist, whose opinion Hill and I respected. Therefore, when he advised us that our conclusion was so important that we ought to be sure that the findings were not due to some quirk of Londoners but could be reproduced in other parts of the country, we had no hesitation in following his advice and arranged to extend the study to include patients in Bristol, Cambridge, Leeds, and Newcastle. Before the second phase of the study could be completed, however, Wynder and Graham reported similar findings in the United States, the findings in the second phase were pointing in the same direction, and there was no reason to wait any longer. With the willing help of the editor, our London results were consequently published in the *British Medical Journal* two months later.

The paper obtained much less publicity than we had expected. Senior clinicians and cancer research workers advising the Ministry of Health were, for the most part, unconvinced of the causal relationship and they advised against publicizing the findings for fear of scaring people. Further evidence of a different type was clearly needed if our conclusion was to be taken seriously and this, we thought, could be obtained by seeking information about people's smoking habits and then following them up to see if the mortality from lung cancer varied, as we predicted it would, with the amount they smoked. For this purpose, Hill suggested that we might seek the help of British doctors, who might be interested in the study and would be easy to follow up, because, for legal reasons, they needed to have their names recorded in an official register of qualified doctors. A pilot study, in which a brief questionnaire was sent to a randomly selected group (which, as it happened included Himsworth, who would never accept that his name had been included by chance) obtained a good response and we decided to press on. With the help of the British Medical Association, we sent out questionnaires to 60,000 doctors with addresses registered in the United Kingdom, and we followed up the 40,000 who replied with the help of the Registrar General (responsible for the country's vital statistics) who notified us of doctors who died. Within $2\frac{1}{2}$ years enough deaths from lung cancer had occurred to show that the mortality was concentrated, as predicted, in the group of cigarette smokers and particularly in those who smoked most heavily. Thus, we felt justified in publishing our preliminary results (*British Medical Journal*, 1954;1:1451–1455). The study was, however, continued for another two years to obtain larger numbers, after which we published what we had intended to be our final result (*British Medical Journal*, 1956;2:1451–1456).

By this time, however, two unexpected things had occurred. The study as planned was concerned only with the effect of smoking on the risk of lung cancer, and we had not expected it to be related to any other important disease, except perhaps to cancers of the upper respiratory and digestive tracts, which we had excluded from the controls in our first study. Patients admitted to hospital for coronary thrombosis (as myocardial infarction was usually called) and patients admitted for chronic bronchitis, as well as patients with many other diseases now known to be related to smoking, were all included in the control series, thus tending to cause an underestimate of the risk from cancer of the lung.

That chronic bronchitis was not then thought to be related to smoking is now almost incomprehensible; but the cough that smokers so often had was called a (benign) "smoker's cough" and patients with chronic bronchitis used to light a cigarette first thing in the morning, in the absence of contrary medical advice, to help clear their chest. When, however, we had finished the second phase of the case-control study and had obtained information about 3,000 patients (half with and half without lung cancer) we noted that smokers more often gave a history of chronic bronchitis than nonsmokers (*British Medical Journal*, 1952;2:1271–1286) and we were interested to see whether, with larger numbers, this would show up in the cohort study. We had also found in the preliminary report of the cohort study that the mortality from coronary thrombosis, for which even then we had records of over 200 deaths, increased progressively with the amount smoked, and this had also been observed by Hammond and Horn in their cohort study of American men and women that had been undertaken by the American Cancer Society specifically (so the senior author told me) to disprove the relationship between smoking and cancer of the lung.

With these two unexpected findings (that smoking was associated with increased risks of chronic bronchitis and coronary thrombosis) and our initial suspicion that it might be associated with some cancers of the upper respiratory and digestive tracts, we decided that the cohort of British doctors should continue to be followed up to provide larger numbers of deaths and show whether any other diseases were also related to the habit. Many doctors were, however, taking note of our findings and giving up smoking and we realized that we needed to approach them again individually, not only to check their vital status, but also to record any change in their habits and obtain some further information about how and what they had smoked in the past. The study was, therefore, continued and is still continuing, with periodic checks on changes in the habits of the survivors, under Richard Peto's guidance, after nearly 50 years.

In retrospect we should, of course, have realized from the beginning that smoking might cause many diseases other than lung cancer, for there were already indications to this effect in the medical literature. The habit was, however, so ingrained and as entrenched in male doctors as in the rest of the population (80% of the middle-aged men being regular smokers at the end of the Second World War) that susceptibility to the idea that it might be an important cause of disease was dulled. The Retrospectoscope was not, however, available to us, nor the ease of searching the medical literature that computers

have made possible, and we started our investigation without any expectation of what we eventually found. The conduct of the work that showed the multiple associations between smoking and disease developed, however, naturally from testing the initial hypothesis. It was, of course, disappointing that Wynder and Graham published their findings before we did ours, but I do not regret it. Himsworth was, I believe, right to advise us not to report our findings until we had shown that they could be duplicated outside London and were, in consequence, confident that they were of general application.