The Health Effects of Smoking

Seminar “How do I lie with statistics”

Marina Walther

Mat.Nr.: 4003856

B.Sc. Computer Science

University of Heidelberg

Seminar “How do I lie with statistics”

supervised by Prof. Dr. Ulrich Köthe

Content

1. Introduction

The health effect of smoking is nowadays commonly accepted as hazardous and dangerous. That is why for example Germany has launched some effective prevention programs, such as “rauch frei”, to prevent especially young adults and children from starting to smoke and being addicted to cigarettes and nicotine. These prevention programs were mainly started right in the beginning of the 21st century. You could ask yourself, why did it take so long to start with public restrictions and prevention programs?

If you have a short look on the history of the global tobacco consumption, you can see that smoking developed to an established and a widely spread factor in the society. In the beginning of the 20th century smoking was more an upper-class thing and for males only. During the first world war smoking turned into a class-independent habit and figures rose to more than 50% and 60% very quickly. Women started to smoke a little later, also as a sign for gender equality and emancipation. Until the 1950s nearly every second woman smoked and 64% of the men. [2] Today, global figures stayed constant with 1.1 Billion people smoking but due to an increase of the world’s population, the proportion of smokers decreased.[1] Especially in Germany the highest decrease of smokers is in the age group of young adults and children. Nevertheless, in Germany 27% of the male adults (older than 18) smoke and 20.8% of the females. [2]

To put it in a nutshell, smoking is still very present in today`s society and according to the WHO one of the top ranked reasons of avoidable death causes. In their latest report from July 2019 the organization assigned 8 million deaths due to tobacco consumption.[1]

This leads to the main emphasis of the following investigation of the health effects of smoking. In the following it will be elaborated and discussed if and how the tobacco industry took influence on the public awareness of smoking and its health hazards. Therefore, the report will be started with a short timeline of important events and developments comparing the public image of cigarette consumption and the inner thoughts of the tobacco industry. Then it will be pointed out, which arguments were held in favour of the pro cigarette image and how they were invalidated or weakened. This will sum up the main points of the “Cigarette Controversy”. Moreover, it will be worked out how the discussion about the health effects of cigarette smoke was dragged on to the topic of secondhand smoke and how the tobacco industry could “recycle” their strategies to create a controversy.

2. Timeline of events

In the 1920s there was the first appearance of an evident linking between smoking and lung cancer through some chemical experiments, for example by Angel Honorio Roffo, an Argentinian chemist. He published from the 1920s to the 1940s various experiments in which he could induce cancer by applying tar from burnt tobacco onto cells. He conducted also some experiments with tar from nicotine free tobacco.[4]

These developments were of course noticed and watched by the tobacco industry. A memorandum from the American Tobacco Company from 1939 gives an evidence for their awareness of Roffo's work. It says, "We have been following Roffo's work for some time, and I feel that it is rather unfortunate that a statement such as his [implicating smoking in cancer] is widely disseminated."[4] This memorandum was just treated internally and did not reach the public in that time. From this and other confidential documents and statements, it can be assumed that senior scientists and executives at tobacco companies knew in the early 1940s about the potential cancer risk of smoking.[4]

A decade later in the 1950s the tobacco companies accepted internally the fact, that smoking caused cancer. This assumption is based on internal confidential statements that reveal their acceptance of evidence. For example, in 1953 Claude Teague, a chemist at R.J. Reynolds Tobacco Company said, "studies of clinical data tend to confirm the relationship between heavy and prolonged tobacco smoking and incidence of cancer of the lung".[4] In 1961 a confidential report to Liggett and Myers Tobacco Company got even more precise. They concluded "[There are] biologically active materials present in cigarette tobacco. These are a) cancer causing; b) cancer promoting; c) poisonous."[5] A report by Dr. Alan Rodgeman, a scientist that worked for R.J. Reynolds referred not only to the evident linking of smoking and lung cancer but also underlined the little opposite evidence of that link. He stated that the amount of evidence accumulated to indict cigarette smoking as health risk was "overwhelming", whereas the evidence challenging such an indictment was "scant".[5]

To manage this upcoming evidence the whole industry started to work together since the early 1950s. The leading industry met to develop a strategy to push the pro cigarette campaign collectively to mislead the public about the dangers of smoking. This was inter alia initiated by the president of the American Tobacco Company in 1953. He invited heads of the leading industry and they decided to employ a public relations firm to develop a response to the smoking and health allegations. Their task is summarized by a memorandum of Hill & Knowlton, the public relations firm which the meeting employed. It said, "We have one essential job – which can be simply said: Stop public panic… There is only one problem – confidence, and how to establish it; public assurance, and how to create it… And most important, how to free millions of Americans from the guilty fear that is going to arise deep in their biological depths – regardless of any pooh-poohing logic – every time they light a cigarette."[5] In the same year the Tobacco Industry Research Committee, short: TIRC, was founded to call for more research, giving the impression of a controversy. The TIRC did not fund any independent research in the first year and by the late 1950s the organization could not be seen as independent from the tobacco industry anymore. Because of that they split off the communications committee of the TIRC and named it Tobacco Institute (TI). The TI was the collective voice of the tobacco industry until 1998 and had the same funding sources, hierarchies and structures than the TIRC. [5]

Publicly, they did not communicate their inner thought and concerns and continued to publish a pro-cigarette campaign. A very famous advertisement from 1954 which is called "A Frank Statement to Cigarette Smokers" was published in major U.S newspapers to question research findings implicating smoking as a cause of cancer. They promised the consumers of cigarettes that their product was safe and vowed to support impartial research to investigate allegations that smoking was harmful to human health.[6] The industry continued to insist that the causal relationship of cigarette smoking and lung cancer is an unproven case, e.g. a spokesman of

the American Tobacco Company stated publicly that "[…] no one has yet proven that lung cancer in any human being is directly traceable to tobacco or its products in any form." He described the evidence of the link between cigarette smoking and lung cancer as "loose talk".[4] The deceptive character of misleading the public opinion is very well expressed by George Weisman in his speech to the National Association of Tobacco Distributors in 1954. The Phillip Morris vice president said, "If we had any thought or knowledge that in any way we are selling a product that was harmful to consumers, we would stop business tomorrow".[6]

On the other hand, the medical world and science got more and more convinced by the fact that smoking is harmful and dangerous for the human health. A leading and very important paper to defend the evident linking of smoking and lung cancer was published in 1958 by Cornfield et al. Their paper about smoking and lung cancer was a pioneering paper in relation to descriptive epidemiological analysis and the investigation of various forms of biases. One of the most cited part of this paper is the formal analysis of risk ratios and the prevalence of an unknown factor "X" that may cause cancer instead of smoking.[7] This mathematical analysis is today known as "Cornfields inequality". The tobacco industry did not publicly accept the facts over the next decades. They got to testify before authorities or made public statements to focus on the lack of evidence of the causal hypothesis and kept on creating a controversy about it. For example, R.J. Reynolds answered a letter to an elementary school teacher in 1968 with "[…] medical science has been unable to establish that smoking has a direct causal link with any human disease."[6] They even testified before the U.S. Congress almost 20 years later in 1982 that "[…] science to date and over a hundred million dollars of our industry`s money indicates that there is no causal link."[6] Internally, they were of course aware of the health hazard of smoking. This comes clear for example with this confidential, internal statement made by Brown and Williamson executives in 1969. It says, "Doubt is our product […] If we are successful in establishing a controversy […], then there is an opportunity to put across the real facts about health and smoking."[8]

First movements in the tight set controversy appeared after some confidential documents of the tobacco industry were given to the University of California in 1994 and leaked shortly after that to public. Just a few month earlier heads of the major U.S. tobacco companies testified before the U.S. Congress their pro-cigarette image and all of them promised to be convinced, that nicotine was not addictive. They answered the corresponding question of the Congress very serious and unequivocal with a clear "No".[9]

However, after the great reveal of these confidential documents, lawsuits amongst others on behalf of the U.S. state governments took place and all lead to restrictions and protection laws, for example the Master Settlement Agreement.[10] With this agreement tobacco companies were restricted with their advertising and marketing practices, as well as to pay various annual payments to the states to compensate for some of the medical costs caused by smoking-related issues. Here you can see some interesting parallels with the figures from the introduction above. Since the 1980s Germans started to smoke less, but effectively prevention programs for young adults or restrictions of smoking in public places were released right after 2000.

3. “The Cigarette Controversy”

3.1. The Tobacco Industry`s strategy

As it was examined above, the tobacco industry tried to create a controversy about the health risks of smoking. Since the 1950s their methods of manipulating research stayed the same over decades. [8]

One of their strategies was to fund research that was in favour of the interest group’s position. There are several purposes of industry-supported research. First, it is very likely that their research draws results and produces findings in favour of the industry itself. Moreover, it is an advantage, that research can be disseminated directly to policy makers and lay press. Of course, with the scientific appearance the credibility of the industry can be increased, and research can provide good public relations by portraying it as philanthropic. [11]

Nevertheless, most of the industry funded research did not undergo any form of scientific peer review and the funding was based on its potential to protect the interests of the companies. Lawyers were highly involved in conduct, design and dissemination of research.[11] Often, it was directly funded by law-firms and the goal of these studies was to protect tobacco companies from litigation and deflect the attention from tobacco as a health hazard. Lawyers selected which projects would be funded and they also decided about the dissemination of the findings. They or consultants prepared an expert testimony for Congressional hearings. Their control reached far more impact, for example it was common that lawyers attended scientific meetings and even reviewed scientific literature. At the tobacco company Brown and Williamson, they even controlled internal scientific reports. All in all, it was ensured that scientific information would be protected from any legal discovery processes.[11] Like it was mentioned in the second paragraph of the report, research organizations and research funding organizations were founded to give the appearance that research was independent from influence of the tobacco industry.[11] Examples here are the Tobacco Industry Research Committee formed by US tobacco companies in 1954 or the Center for Indoor Air Research formed by three major tobacco companies, Philipp Morris, R.J. Reynolds and Lorillard Corporation in 1988. These research committees were often criticised because of the obvious relations and dependencies to the tobacco companies. The Center for Indoor Air Research (CIAR) was even disbanded by the Master Settlement Agreement in 1998 but the industry reacted by renaming these institutions or forming new ones with the same structures and hierarchies. That is why for example the CIAR was replaced after its disbandment by Phillip Morris in 2000. He formed a new research program with similar structures than the CIAR. [12]

In addition to the highly involved funding process the industry also published a lot of research in favour of their position. They published in symposiums, books, journal articles etc. and then cited its industry-funded, non-peer-reviewed publications in scientific and policy arenas.[12] Especially the symposium proceedings delivered a lot of seemingly scientific valuable results because from an outer view these symposia presentations look like peer-reviewed journal articles. The tobacco industry sponsored numerous symposia especially on the issue secondhand smoke and payed for scientific consultants to organize and attend those. From 1965 to 1993 eleven symposia lead to publications. From these 11 were six published as special issues in medical journals and five lead to independent books. None of these publications were peer reviewed but were often cited as if they had been. Moreover, the positions taken by symposium presenters were described as he consensus of a gathering “of leading experts from around the world” who disagreed with the public literature on secondhand smoke.[13]

Not only special proceedings lead to a disproportionate influence on policy. The quality of tobacco industry funded publications was often rather poor. Studies have examined different characteristics that might be associated with the quality of a scientific paper. This includes the

disclosure of sources of research sponsorship, the article conclusion, the article topic and study design. [13]

Furthermore, the publications did not have any industrial co-authors, but one single scientific author presented as the sole author of the paper. The financial or industrial dependencies were often not clearly disclosed or even not disclosed at all. The common way of mentioning an industrial dependency was to set a little and inconspicuous note of thanks at the very end of the publication. [14]

In addition to that, there was a simple strategy to get their results into policy arenas. First industry favourable findings, data or interpretations of risk were disseminated in the lay press. This was important to stimulate a controversy in the lay print media about the health effects of smoking and leave the readers with the impression of an ongoing character of the controversy. Then, results, data and interpretations of risk were presented directly to policy makers. The public participation was of course important for the acceptance of the findings and the shape of the final risk assessment or regulation.[15]

Besides the industry's attempts to create scientific research by themselves they also had an interest in suppressing and minimizing unfavourable findings and scientists working for the opposite site of argumentation. There were legitimate means included, such as letters to the editor of scientific journals and editorials. But on the other side there were also less legitimate means, such as attacks towards the integrity of researchers and lawsuits to obtain data that are then reanalysed. In some cases, industry consultants were paid to criticize independent research on tobacco and also secondhand smoke. [15]

3.2. Main aspects of the controversy

In the following paragraph it is going to be expounded which arguments in particular were held against the hazardous health effect of tobacco smoke and which arguments support the causal linking of smoking and diseases such as lung cancer. The main part of the so called "Cigarette Controversy" took place in the middle of last century approximately from the 1940s on. Of course, this needs to be taken into consideration when evaluating the methods of disproving or weakening the arguments of the non-causal party. It is important to note that some arguments that were brought up against the causal relationship could be easily disproven with facts, but some arguments were absolutely valid aspects of criticism and worth to discuss at that time and still nowadays.

The confrontation is mainly based on the pioneering paper of Cornfield and his colleagues about "Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions" published in 1958. It was one of the most accurate papers discussing the causal relationship between smoking and lung cancer by pointing out recorded facts on the one hand. On the other hand, they defined more precisely inadequacies of information and open points of the discussion where further research was necessary. They cited 21 retrospective studies, all independent groups from eight different countries and three prospective studies, also independent groups from two different countries. All of them had a similar conclusion in favour of the hypothesis that smoking causes lung cancer. They summarized that the information already available was sufficient for planning and activating public health measures. Nowadays, this paper is categorized as a part of the origin of causal reasoning in today`s epidemiology and the root of some important epidemiological practices as they are done today. [16]

3.2.1. Descriptive epidemiological arguments

Effect of Aging:

Some critics were since the population at that time already had an aging development. It was questioned if the aging population could have caused a higher death rate of lung cancer. Cornfield et al. responded correctly with the fact that this issue was investigated by age

adjusted rates. For example, a study by Dunn examined the U.S. population used age adjusted rates and concluded that only 1/6 of the deaths was due to the effect of aging. There were also similar findings not only for the U.S. but also for English and Welsh studies. Moreover, this counterargument can be disproven with the fact that age-specific lung cancer rates are not affected by an aging population but still show a significant increase.[17] The methods of disproving the counterargument of an aging population are very similar to today`s practices. It is a common way to use age standardization either based on a standard population or based on standard rates to validate the study's population data. Furthermore, methods of data stratification are commonly used to control static variable and confounders such as the effect of an aging population.

Improvement in Diagnostic Factors:

Another counterargument referred to the general improvement of diagnostic measures and a more complete medical reporting. Concrete critics had seen the responsibility of increasing lung cancer rates in better methods of the diagnosis and recognition of the disease and improved medical documentation and reporting. Regarding necropsy data they claimed that there had been no real increase of the corresponding death rates and the effect was just due to statistical classification. Cornfield et al. tried to weaken this criticism with some "special features" of the lung cancer rate increase. For example, a rising ratio of male to female deaths could not be explained by the improvements of diagnostic factors or more complete medical reporting. They also mentioned that successively younger cohorts are affected by the rising appearance of lung cancer. The main arguments against that criticism are first a careful study published by Gilliam. He took possible incorrect certified lung cancer deaths into consideration and counted them as tuberculosis cases and came to lower bounds of an increase of lung cancer deaths. There was still a fourfold higher risk of dying of lung cancer for males instead of a 26-fold one and a 30% higher risk for females instead of a sevenfold one.[17,18] Secondly, they referred to the Copenhagen Tuberculosis Station Data, a registry of tuberculosis incidents where procedures of diagnosis remained unchanged from 1941 to 1950. The data from that registry helped to measure the cofounding effect of improved diagnostic methods because they could compare the data to the Danish cancer registry. The prevalence rate among male examinees increased by the rate recorded by the Danish cancer registry whereas the time related data from the Copenhagen Tuberculosis Station did not show any increase of incidents. This was an important aspect of their argumentation.[18] In addition to that, necropsy data even helped to weaken the criticism instead of strengthening it. Cornfield emphasized that necropsy data had mainly a high-quality diagnosis but a lower quality report in contrast to mortality data, which is mainly based on death certificates. These have the attributes of a low-quality diagnosis but a detailed report. The fact that both of these sources agree with each other would mean that both approaches lead to the same inferences. From today's point of view their way of argumentation was very broad but still not that sharp as it was regarding the argument of an aging population.[18] The improvements in diagnosis and medical reporting are still hard to measure and are currently tried to be examined for example with studies on the effectiveness of medical screening programs. These studies are fighting with various forms of biases and therefor seen critically.

Socioeconomic status of smokers

Critics of the causal relation between smoking and lung cancer also see a counterargument in inadequacies of studies concerning the socioeconomic status of the population group. They stated that smoking was not a class dependent appearance but there is a significant higher lung cancer rate among lower income classes. How could a causal hypothesis explain the effect of lower income classes having higher amount of lung cancer incidents. They concluded that there must be other factors that contribute more to the causation of lung cancer. Cornfield and his colleagues could not disprove the argumentation properly, but they tried to weaken the counterargument with various informal thoughts. They underlined that smoke is not the sole cause for cancer, therefore it could not be applicable as the sole explanation for increasing lung cancer rates. Furthermore, the population exposed to other established agents, such as

industrial carcinogens and toxics, is too small to draw further conclusions for the whole population. They also denied that the effects of the socioeconomic status were not big enough to have a significant impact compared to the impacted noted for smoking history.[19] You can see, that Cornfield's attempt to weaken the criticism was very loose and not that sharp. Today confounding is still a big problem in epidemiological data analysis but there are various approved ways to deal with it in a more correct and formal way. First of all, studies are designed in the aspect of controlling static variables and confounders, for example with the restriction of participants, matchings strategies and randomization to choose a legitimate study population group. Afterwards with the analysis of the results there are methods of adjustment, such as standardization and stratification, and multiple regression to clean the findings.

There was further criticism with similar ways of creating a debate about descriptive factors, for example the sex difference [20], which means that the lung cancer rate is higher among males than females, or the urban-rural difference. Critics tried to emphasize that a causal hypothesis seemed illogical with a higher lung cancer rate among people living in cities than people living in rural areas whereas the smoking habit would be similar. These arguments against the causal relationship of smoking and lung cancer were rather vague and easy to disprove with facts about the specific population. They could easily turn into a supportive argument in favour of the causal relationship of smoking and lung cancer. [18,20]

3.2.2 Special Population Groups

The Haag Hanmer-study

With the study by Haag and Hanmer investigating nine processing plants of the American Tobacco Company, the tobacco company published their own scientific research based on the assumption that smoking did not cause cancer. They stated that the study population had an above average proportion of smokers but no higher mortality for respiratory cancer or coronary diseases compared to the whole population of Virginia and North Carolina. The study concluded that the existence of such population made it evident that cigarette smoking per se is not necessarily or invariably associated with a higher risk of lung cancer or cardiovascular diseases or with diminished longevity. [19]

The findings of the Haag Hanmer-study are based on wrong assumption and the study was not designed to draw a wide conclusion related to the whole population. One major problem was the rather small sample group. The investigation was based on nine processing plants and therefore the number of participants of the study was too small to draw conclusions on respiratory cancer affecting a whole population. Also, the study population consisted of Employees only. Employees do not represent the general population as the working population is in general healthier as the total population. The comparison in the study between the study population and the whole population of North Carolina and Virginia is because of that inapplicable. Analysing the weaknesses of the study there are far more problems occurring. One major problem is that there is no further information about the study population provided. The publication lacks important data for the smoker/non- smoker proportion and the selection of employees. [21] You would also need the information if there were only the currently working people part of the study or also the retired ones or how they counted the former employees who had already died.

All in all, there are too many inadequacies in this study to draw a conclusion stated to be a valid evidence for the whole population and the cigarette smoking as a health hazard debate.

3.2.3 The Selection of study groups

There was a lot of criticism concerning the studies' design, in particular the selection of study groups in retrospective studies. The retrospective method from the time the paper was published refers to the current case-control model. The criticism was an absolute valid discussion about various forms of selection biases, as it would be today titled. They pointed out important weaknesses of case control studies that have nowadays still an important impact

on study-design. More specifically, critics held against the retrospective method that almost all the studies had a control group without lung cancer incidents. The control group should represent the general population which of course includes lung cancer patients. Another point of criticism in relation to the selection was that the study population usually consisted of hospitalized patients only. Patients in hospitals would not represent the smoking habits of the general population. [23] Moreover, critics estimated that people with two complaints are more likely to get hospitalized than people with only one complaint. This would distort the health hazardous effect of smoking with critical patients and comorbidities.[22] They did not only criticize the methods of choosing the studies' samples they also argued that a survey at a given instant of time would lead to misleading results.[22] This meant that a case control study takes usually place in a strictly regulated time span of some weeks or some years and is not designed as a long-term survey. Cornfield and his colleagues appreciated the discussion about the weaknesses of the retrospective method, but they pointed out that all the cited studies had a similar conclusion and often when the criticism is looked at in particular it only was applicable to some of the studies.[22] Their counterargument was most of the time that various different study designs came to the same conclusion and therefor the weaknesses of the methods could not be sufficient to minimize the strengthen of the causal hypothesis itself. Furthermore, a higher mortality rate in later periods of the studies was observed. Cornfield et al. interpreted this effect as a direction of the bias that would underestimate the smoking-health-association. [24]

Of course, the discussion about the different forms of selection bias was held in a very informal way but the criticism about the method itself was valid and worth to debate about. Nowadays the selection bias is still a big problem in medical studies and with the randomized and blinded design of case control studies it is taken a big effort to control biases like this.

3.2.4. Accuracy of information

Besides the criticism on the retrospective method, there was also a debate about the weaknesses of the prospective method. First of all, it is important to notice that at the time the paper was published prospective studies correspond to the model of cohort studies we know today. Critics at that time pointed out that a certain error occurred with the ascertainment of smoking habits and the diagnosis of the disease during the progress of the study. [24] The argued error was especially based on the fact that patients may be biased responding to surveys about their smoking habits and physicians my also be biased diagnosing people too early with the disease to keep the progress of the study promising. [25] The criticism was sure valid and important to debate about as this form of bias is nowadays well known as the information bias. Against a serious impact of the bias in relation to the collection of data about smoking habits, Cornfield et al. brought up that survey results yielded like figures on tobacco production or taxation. [25] This effect was underlined by the reproducibility of the observation several years later. Moreover, they cited a study that had investigated the reliability and accuracy of the patient's or the patient family's replies. The study showed on the one hand that the answers of the participants had not been accurate, but few heavy smokers were classified as light smokers.[25] One could ask yourself if a response to the criticism on a possible information bias like this is sufficient to accept the bias as weak enough to ignore it. By that time such argumentation was certainly the right direction but still very informal and vague. The danger of a bias concerning the diagnosis of lung cancer throughout the study was declared as well founded but just as a major problem for the prospective method. In addition to that Cornfield and his colleagues held against the criticism that since all deaths had got to be included, there was still the fact that the studies had found a higher death rate for cigarette smokers. The death rate had even increased the stronger the smoking habits got. If now the relation between their death and their smoking habits had been left out there were only two logical conclusions. On the one hand the association between cigarette smoking and lung cancer was somewhat overestimated but on the other hand this would lead to a high underestimation of another association of cigarette smoking and some other death leading disease.[24] Like the debate about the selection bias, the criticism on the accuracy of information was well founded. Today in order to control the information bias, the design of the

study needs to be well thought and discussed, to prevent the data collection to be biased by the interviewer or the interviewed participant. Therefore, there are again blinding methods indispensable.

3.3 Interpretation of results

When it comes to the interpretation of the findings of studies there were different approaches to explain an increased death rate or lung cancer risk ratio among smokers. The hypothesis Cornfield and his colleagues stood up for is called the causal hypothesis. This means that the explanation of the increased rates would be as simple as smoking causes lung cancer. On the other side, often pro-cigarette parties, claimed that there could be a still unknown common cause, an unidentified factor X, that would cause both smoking and lung cancer. This would include that there is no further direct association between smoking and lung cancer. Fisher, a famous statistician of that time, also mentioned that there was the possibility of a third interpretation. He argued that for tor the sake of logical completeness lung cancer causing smoking could be discussed. Cornfield et al. did not investigate the third interpretation any further. [26]

There were some specific arguments that disturb the logic of the constitutional hypothesis. One major breaking point was that the mortality in lung cancer had increased continuously from 1900 to the 1950s considerably more for males than for females. This would assume that a possible environmental change which would be responsible for that unknown constitutional factor had been more sensitive for males than for females. One could also suggest the higher increase for males as a result of a sex-linked mutation. All in all, it would require an environmental agent other than tobacco that males being exposed to it considerably more than females. Exposure would then be as highly correlated to lung cancer than tobacco and it would also be as highly correlated with cigarette consumption. [27] Cornfield et al. mentioned not only logical inadequacies with the sex gradient of the increased lung cancer rate but also drew attention to the experimental induction of cancer with tobacco tar. You would need to assume a simple coincidence with the application of tar onto bronchogenic epithelium of mice and rats and the induction of cancer.[28] Moreover, studies had shown that different types of tobacco consumption, heavy / light smokers or inhaling / non-inhaling smokers, lead to different types and stages of cancer.[28] Studies had also shown that discontinued smokers had a lower mortality than lifelong smokers. Here the constitutional hypothesis lacks again explanation. They questioned if it was thinkable that the possible constitutional factor, e.g. a gene, which predisposes lung cancer and the craving for cigarettes could be a variable and changing characteristic of people. [28,29]

The most popular yet unidentified constitutional factor that was suggested was a genetic one. Fisher proposed a gene together with the results of a study that investigated the smoking habit of monozygotic and dizygotic twins. He observed that 51 monozygotic twins had had more similar smoking habits in comparison to other 33 dizygotic twins. Of course, the observation leads to the conclusion that there could be a gene induce cigarette smoking and the craving for tobacco. But in the end, there could not be drawn a conclusion with a certain statement about the association between smoking habits and lung cancer. [29]

The most important counterargument against the constitutional hypothesis is Cornfield et al.'s investigation about the possible constitutional factor "X". They stated that with a nine-fold greater lung cancer risk among cigarette smokers, the prevalence of the factor "X" among smokers must be at least nine times greater as the prevalence of that factor "X" among non-smokers.[29] This rather informal description of Cornfield's inequality was mathematically proved in the appendix of the paper. In the following the derivation of the stated inequality will be expounded. [based on: 30,31]

Cornfield's inequality:

Let D be the presence of the disease, B the presence of the causal agent and A the presence of the unobserved agent. Let $\bar{B}$ be the absence of the causal agent and $\bar{A}$ the absence of the unobserved agent.

Let further notation be $f_1 = P(A|B)$ and $f_0 = P(A|\bar{B})$ to improve readability.

Now consider the risk ratios R_o, the risk ratio for the observed agent, and R_u the risk ratio for the unobserved agent. The risk ratios are computed with the corresponding prevalence P.

$$R_0 = \frac{P(D|B)}{P(D|\bar{B})}$$

$$R_u = \frac{P(D|A)}{P(D|\bar{A})}$$

Now assuming D to be independent from B with a given A, these equations follow:

$$p = P(D|\bar{A}, B) = P(D|\bar{A}, \bar{B}) = P(D|\bar{A}) \tag{i}$$

$$pR_u = P(D|A, B) = P(D|A, \bar{B}) = P(D|A)$$

Having a look at R_o, it applies:

$$R_0 = \frac{P(D|B)}{P(D|\bar{B})} = \frac{P(D, A|B) + P(D, \bar{A}|B)}{P(D, A|\bar{B}) + P(D, \bar{A}|\bar{B})}$$

$$= \frac{P(D|A, B)P(A|B) + P(D|\bar{A}, B)P(\bar{A}|B)}{P(D|A, \bar{B})P(A|\bar{B}) + P(D|\bar{A}, \bar{B})P(\bar{A}|\bar{B})}$$

$$\overset{\text{(i)}}{=} \frac{pR_u f_1 + p(1 - f_1)}{pR_u f_0 + p(1 - f_0)} = \frac{R_u f_1 + (1 - f_1)}{R_u f_0 + (1 - f_0)} \tag{ii}$$

(ii) maximizes if you consider f_1=1 and f_0=0 and assume $R_u \geq 1$ and fixed. Then it implies

$$\boxed{R_0 \leq R_u} \tag{iii}$$

(ii) also maximizes if you do not assume R_u to be fixed but let R_u go towards infinity. It follows

$$R_0 \leq \frac{f_1}{f_0} \qquad \text{(iv)}$$

(iii) and (iv) both prove the based statement and are known as Cornfield's inequality.

3.4. Causality and why it is critical to assume it

At the very first glance, it might seem obvious that the causal hypothesis is the "correct" interpretation. Of course, the cigarette controversy itself was not only a clean and fair discussion about the pro and counterarguments of the causal relationship between smoking and lung cancer, but some criticism on the methods to validate the causal interpretation was very well founded. Finally, one is wondering, why is it so hard to prove causality in a way that criticism cannot unsettle a solid interpretation of scientific findings.

The causal hypothesis is from a critic's point of view solely based on "statistics" and lacks experimental evidence. Experiments in which tar was applied onto cells or experiments with laboratory animals showed a strong association between tobacco smoke or tar and the development of cancer. Critics said that all these experiments were in the end not applicable to the human case. To shut down these critics there are experiments on human subjects over a long period like 30 to 60 years necessary. These kinds of experiments are of course ethnically inconceivable.

Another aspect that disturbs the proof of causality is that there are still non-smokers who suffer from lung cancer and smokers who enjoy a cancer free live. Of course, cigarette smoking is not the sole reason people develop lung cancer, but these cases do not support the convincing of people who have doubts about the causal relationship.

To put it in a nutshell all methods and studies have their weak points, but all evidence together seems to be in favour of the causal hypothesis. Therefore, the more correct formulation of causality would be that all evidence together makes the causal hypothesis seem much more likely to be the correct interpretation. [32]

4. Recycling strategies: Secondhand tobacco smoke

The above investigated controversy was all about the health effect of cigarette smoking itself. In the following there will be a short investigation along an example study of the issue secondhand tobacco smoke and the health effects for non-smokers in particular. It is interesting to note that the tobacco industry could keep their strategies developed to create a pro-cigarette image in the 1950s to take influence on the public opinion about the health effects of secondhand tobacco smoke.

4.1. The 16 cities study and its design

The so called 16 cities study will be used as an example of how the tobacco industry manipulated scientific research to produce findings in favour of their interests.

The study was designed, executed and supervised by R.J. Reynolds Tobacco Company scientists. The goal of the study was to investigate the secondhand smoke exposure at workplaces compared to the exposure at the worker's homes. The study was published in 1996 by Jenkins who was displayed as the sole author of the publication. The publication did not disclose the full involvement of R.J. Reynolds, although the tobacco company had major financial and content impact on the study. [33]

To anticipate the result of the study, it was concluded that home secondhand smoke exposure was four times greater than workplace secondhand smoke exposure. Therefore, the suggestion was that no legal regulations would be necessary because of the insignificant impact of workplace secondhand smoke exposure.[34]

Analysing the interior statements and documents concerning the 16 cities study it can be subsumed that the purpose of the study was to affect policy making and not advancing scientific knowledge.

The participants were selected from 16 cities across the U.S. and had to be non-smokers. The study was in that way innovative and scientifically worthy as there were no detailed data sets about secondhand smoke exposure at that time. They used innovative methods to create a useful and detailed data set which were laborious and costly. Every participant had a sampling pump during work that measured the secondhand smoke exposure with chemical markers. Then the participants had a separate sampling pump for their time staying at home. The whole analysis of the data received from the sampling pumps was based on measurable chemical markers. Besides the data from the pumps every participant had to write a diary about the number of cigarettes being smoked in their environment. The subjects had to count the number of cigarettes being smoked within 100 ft of themselves every hour during the air sampling process. [35]

4.2. Published data analysis

Reviewing the data analysis published with the original study some inaccuracies and mistakes can be observed.

4.2.1 Omitted data

First of all, not all of the data collected during the sampling process of the study was published with the original study. They omitted data on what was written in the diaries of the subjects. After asking the publisher why this seemingly valuable data was omitted the response was short and vague. They stated that the raw data of the diaries were just self-reported observations and therefore its value unreliable. What the data really showed was a significant association between the number of cigarettes being smoked in the ultimate surroundings of

the subjects and the concentration of the secondhand smoke components being measured by the sampling pumps.[35] It is very interesting to note that in 1999, just a few years after the original study was published, the author of the 16 cities study used the same data to co-author a publication that supports the association implied by the diary-data. This supports the high degree of influence that R.J. Reynolds took to control the outcome of the study. [35,34]

Furthermore, the data analysis that was done with the collected data had some misleading assumptions that were not correctly disclosed.

4.2.2 Definition of a “smoking workplace”

The 16 cities study classified smoking workplaces in a binary way. In the publication it was nowhere clarified that the study’s definition of a smoking workplace included workplaces where smoking was already restricted to designated areas. The partition of workplaces categorized as a smoking environment although it was only permitted in designated areas was greater than 68%. The proportion of somehow misclassified smoking workplaces has a significant impact on the conclusion drawn from the data. [34] Barnes et al. computed with the same data that the median of the nicotine concentrations in non-restricted smoking workplaces is 6.5 times higher than in smoking restricted workplaces. This estimation is confirmed by another study by Hammond who computed a 6.6 times higher median for the non-restricted smoking workplaces. [34]

Originally the study was intended to have 50% of its sample data from people who worked in smoking workplaces, but the final sample conditions showed that only 10% of the participants worked in places where smoking was not restricted.

In addition to the above discussed inadequacies there is still the omitted information about the personal observation of the subjects needed to be aware of. Through the analysis of the diary data the published conclusion gets more disturbed. Less than a third of the subjects who worked in surroundings where smoking was restricted observed smoking at any time. Also, those who did observe smoking in restricted workplaces reported significantly fewer cigarettes being smoked than those who work in non-restricted smoking workplaces.[34]

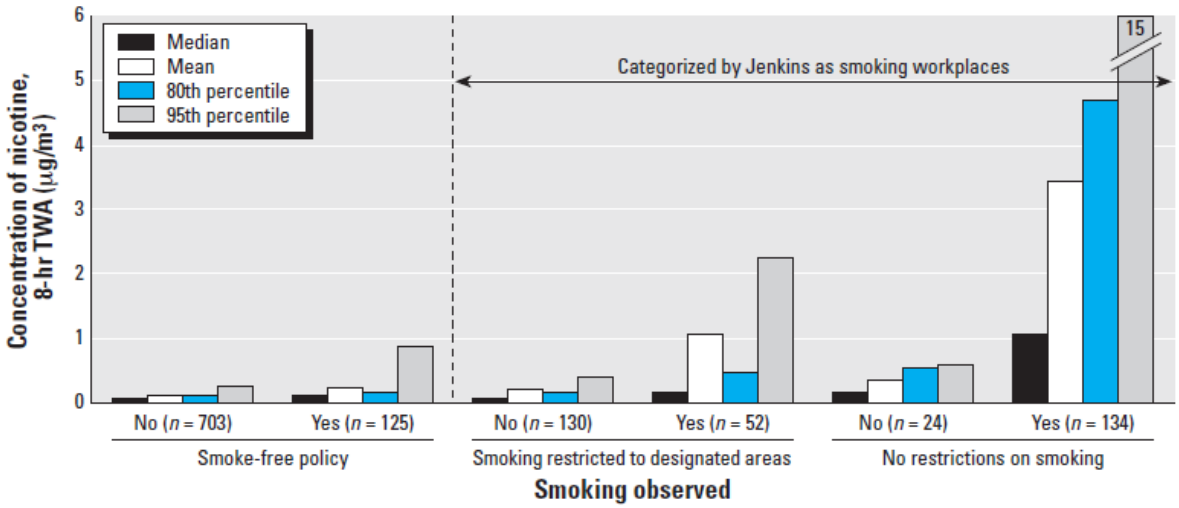

Fig. 1: Misleading classification of smoking workplaces

4.2.3 Inappropriate comparisons of cells

The conclusion drawn by the originally published study was based on a misleading comparison of cells.[34] The based four-field table consists of unstratified data that was globally grouped. This leads to a comparison of people who work with smokers and people who live with

smokers. Within this comparison the cell “smoking home and smoking workplace” is not distinctly assigned. Moreover, the comparison is not suitable for the how workplace exposure adds to the total exposure for individuals. This question would effectively answer the concerns about secondhand smoke exposure at places, where legal restrictions could be applied and then be a benefit for people’s health. [36]

4.3. Reanalysing the data

If you would apply a different approach on analysing the data sets from the 16 cities study, you would come to a different conclusion. The data set collected by Jenkins is suitable and helpful to investigate further questions on the secondhand smoke exposure at workplaces.

First you could stratify the data by the home smoking status of the subjects, then the relative exposures are reported of those who work with smokers and those who work in a smoke-free environment. After stratifying the data that way one can see that the exposure gets added up. Subjects who lived together with smokers and then worked together with smokers doubled their exposures compared to those who work in non-smoker environments. The exposure for subjects who live in a smoke-free home but work at a smoking workplace is even more than ten times higher compared to those who work and live in a non-smoking environment. [36]

The different analysis of the dataset would conclude a serious minimization of secondhand smoke exposure if smoking would be prohibited at workplaces. This would lead to a significant health benefit for non-smokers. [36]

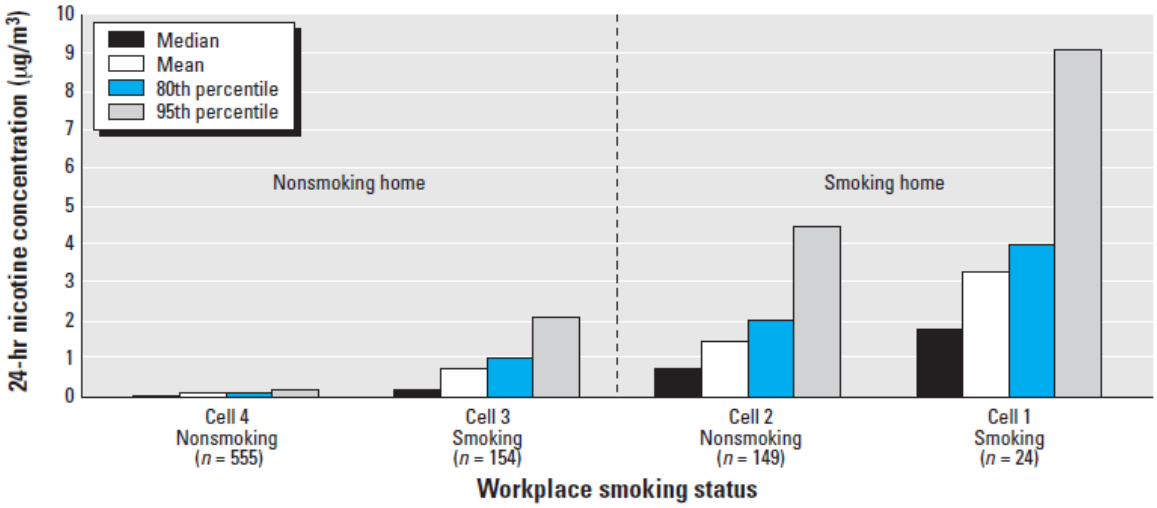

Fig. 2: Stratified data leads to a different analysis

5. Conclusion

The main goal of this report was to investigate on the one hand the methods of manipulating scientific research by the tobacco industry, but this was not the sole intention. The health effects of smoking seem to be an easy, indisputable question to answer. Therefore, it was also important to reflect a debate about an issue which is from today`s point of view easy or obvious to reach a conclusion for. Of course, industry funded research did not always pursue the intention of improving scientific knowledge about the health effects of smoking but there was also criticism that was well founded and needed to be discussed. Especially the publication by Cornfield et al. had a pioneering character of discussing several aspects which are still very important in today's analytical epidemiology. They also debated about weaknesses of various study models, nowadays described as the information and selection bias that are still the most important forms of biases to control in medical studies.

It is also obvious that today scientists still have got to be cautious when participating in industry funded projects or using and reviewing results from it. Nevertheless, the scientific world is reliant on funding and support from the industry.

That is why critics demand more transparency. They demand a full disclosure of financial dependencies and a full disclosure of industrial involvements to ensure the integrity of scientific research. [37, 38]

In the end, a well formulated saying is a good way to summarize not only the relationship between industry and scientific progress but is also a good advice for safe and serious research. It says,

"Do not trust a study you did not manipulate yourself.".

Bibliography

4: Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions

5: Barnes, R.L, Hammond, S.K, Glantz, S.A. 2006. The Tobacco Industry's Role in the 16 Cities Study of Secondhand Tobacco Smoke: Do the Data Support the Stated Conclusions? Environmental Health Perspectives 114(12): 1890-1897.

6: Bero, L.A. 2005. Tobacco Industry Manipulation of Research. Public Health Reports 120: 200-208

7: Cummings, K.M., Brown, A., O'Connor, R., 2007. The Cigarette Controversy. CEBP Focus: Nicotine and Tobacco-Control Research 16(6): 1070-1076

8: Hong, M-K., Bero, L.A. 2002. How the tobacco industry responded to an influential study of the health effects of secondhand smoke. BMJ 2002; 325:1413

9: Barnes, D.E., Bero, L.A. 1998. Why Review Articles on the Health Effects of Passive Smoking Reach Different Conclusions. JAMA 279(19): 1566-1570

10: Greenhouse, J. 2009. Commentary: Cornflied, Epidemiology and Causality

11: Zwahlen, M. 2009. Commentary: Cornfield on cigarette smoking and lung cancer and how to assess causality

12: Smith, G.D. 2009. Smoking and lung cancer: causality, Cornfield and an early observational meta-analysis

13: Cox, D.R. 2009. Commentary: Smoking and lung cancer: reflections on a pioneering paper

14: Vandenbroucke, J.P. 2009 Commentary: "Smoking and lung cancer" – the embryogenesis of modern epidemiology

Sources and Citations

1. https://www.tagesschau.de/ausland/rauchen-tod-who-101.html
2. https://www.dkfz.de/de/tabakkontrolle/download/Publikationen/sonstVeroeffentlichungen/Tabakatlas-2015-final-web-sp-small.pdf
3. https://www.bundesgesundheitsministerium.de/service/begriffe-von-a-z/r/rauchen.html
4. Cummings, K.M., Brown, A., O'Connor, R., 2007. The Cigarette Controversy. CEBP Focus: Nicotine and Tobacco-Control Research 16(6): 1071
5. Cummings, K.M., Brown, A., O'Connor, R., 2007. The Cigarette Controversy. CEBP Focus: Nicotine and Tobacco-Control Research 16(6): 1072
6. Cummings, K.M., Brown, A., O'Connor, R., 2007. The Cigarette Controversy. CEBP Focus: Nicotine and Tobacco-Control Research 16(6): 1073
7. Greenhouse, J. 2009. Commentary: Cornflied, Epidemiology and Causality
8. Bero, L.A. 2005. Tobacco Industry Manipulation of Research. Public Health Reports 120: 200
9. Cummings, K.M., Brown, A., O'Connor, R., 2007. The Cigarette Controversy. CEBP Focus: Nicotine and Tobacco-Control Research 16(6): 1070
10. Cummings, K.M., Brown, A., O'Connor, R., 2007. The Cigarette Controversy. CEBP Focus: Nicotine and Tobacco-Control Research 16(6): 1074
11. Bero, L.A. 2005. Tobacco Industry Manipulation of Research. Public Health Reports 120: 201
12. Bero, L.A. 2005. Tobacco Industry Manipulation of Research. Public Health Reports 120: 202
13. Bero, L.A. 2005. Tobacco Industry Manipulation of Research. Public Health Reports 120: 203
14. Hong, M-K., Bero, L.A. 2002. How the tobacco industry responded to an influential study of the health effects of secondhand smoke. BMJ 2002; 325:1413, 5
15. Bero, L.A. 2005. Tobacco Industry Manipulation of Research. Public Health Reports 120: 204
16. Zwahlen, M. 2009. Commentary: Cornfield on cigarette smoking and lung cancer and how to assess causality
17. Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, 176
18. Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, 177
19. Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, 179

20. Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, 178

21. Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, 180

22. Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, 181

23. Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, 182

24. Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, 183

25. Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, 184

26. Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, 190

27. Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, 191

28. Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, 192

29. Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, 193

30. Greenhouse, J. 2009. Commentary: Cornflied, Epidemiology and Causality, 1200

31. Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, 198-199

32. Cornfield, J., Haenszel, W., Hammond, E.C., Lilienfeld, A.M., Shimkin, M.B., Wynder, E.L. 1959. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions, 196-197

33. Barnes, R.L, Hammond, S.K, Glantz, S.A. 2006. The Tobacco Industry's Role in the 16 Cities Study of Secondhand Tobacco Smoke: Do the Data Support the Stated Conclusions? Environmental Health Perspectives 114(12): 1891

34. Barnes, R.L, Hammond, S.K, Glantz, S.A. 2006. The Tobacco Industry's Role in the 16 Cities Study of Secondhand Tobacco Smoke: Do the Data Support the Stated Conclusions? Environmental Health Perspectives 114(12): 1894

35. Barnes, R.L, Hammond, S.K, Glantz, S.A. 2006. The Tobacco Industry's Role in the 16 Cities Study of Secondhand Tobacco Smoke: Do the Data Support the Stated Conclusions? Environmental Health Perspectives 114(12): 1893

36. Barnes, R.L, Hammond, S.K, Glantz, S.A. 2006. The Tobacco Industry's Role in the 16 Cities Study of Secondhand Tobacco Smoke: Do the Data Support the Stated Conclusions? Environmental Health Perspectives 114(12): 1895

37. Barnes, R.L, Hammond, S.K, Glantz, S.A. 2006. The Tobacco Industry's Role in the 16 Cities Study of Secondhand Tobacco Smoke: Do the Data Support the Stated Conclusions? Environmental Health Perspectives 114(12): 1896

38. Bero, L.A. 2005. Tobacco Industry Manipulation of Research. Public Health Reports 120: 206

Image Sources

Fig.1: Barnes L. Richard, Hammond S. Katharine, Glantz A. Stanton 2006: The Tobacco Industry‘s Role in the 16 Cities Study of Secondhand Tobacco Smoke: Do the Data Support the Stated Conclusions? Environmental Health Perspectives 12: 1894

Fig.2: Barnes L. Richard, Hammond S. Katharine, Glantz A. Stanton 2006: The Tobacco Industry‘s Role in the 16 Cities Study of Secondhand Tobacco Smoke: Do the Data Support the Stated Conclusions? Environmental Health Perspectives 12: 1895