The Association between Smoking and Alzheimer’s Disease: Effects of Study Design and Bias

Walter A. Kukull

In epidemiologic studies, unrecognized bias can contribute to observed results, causing them to be inaccurate. Analytic study designs, such as the case-control and cohort designs, each carry potential for specific forms of bias. The cohort design is not susceptible to many forms of bias that are experienced by case-control studies. A consistent “protective” effect of smoking on Alzheimer’s disease was documented by many case-control studies. However, the potential effect of biases cannot be separated from the results. Cohort studies now show that smoking may either be unrelated to Alzheimer’s disease onset or possibly generate a modest increased risk. In this review the results and comparisons of various studies and potential biases are discussed.

Key Words: Dementia, Alzheimer’s disease, epidemiology, smoking, bias, epidemiologic methods

Introduction

Scientific evidence based on randomized, controlled trials (RCTs) is generally regarded as a high standard of proof of association between a factor and an outcome. In RCTs the investigator randomly assigns comparable patients to one of several treatments. Because valid randomization equalizes the distribution of known and unknown factors related to outcome, differences between the groups can be attributed to the treatment that the investigator applies. However, there are many situations where it is either unfeasible or unethical to apply this research design.

Observational studies to assess the effect of a particular exposure on disease onset are alternatives to RCTs. Careful attention to detail in patient selection, information gathering, and exposure assessment are critical to the validity of observational studies. Studies that assume the temporal relationship between exposure and disease add to our ability to determine whether associations may be causal. Two common observational, analytic study designs used in epidemiology (Gordis 1996; MacMahon and Trichopoulous 1996; Rothman and Greenland 1998) are the “case-control” design (or case-referent, case-comparison, etc.) and the “cohort” study design (prospective, concurrent, or historical). In the case-control studies, persons with a specific disease (cases) are assembled along with a comparison group of persons without the disease of interest. History of “exposure” before disease onset is then determined. In cohort studies, exposures are determined for a group without the disease of interest and the incidence of disease is then followed over time. The structure is similar to that of an RCT except that the investigator has no control over “exposure group” assignment.

Between the 1980s and mid-1990s a number of case-control studies were conducted to study the effect of smoking on the onset of Alzheimer’s disease (AD). Lee (1994) produced a review and overall analysis of the potential effect of smoking on AD. He carefully evaluated potential study differences that could have had an effect on results, and then calculated a summary measure of effect based on the reported odds ratios, study size, and confidence intervals (CIs). By definition, the odds ratio is the odds of smoking among the cases divided by the odds of smoking among the control subjects. The odds ratio is interpreted as an estimate of the relative risk; therefore an odds ratio of 2.0 would indicate that smoking was twice as common among cases or that smoking was associated with a twofold increased risk of AD; an odds ratio of 0.5 would imply that smokers enjoyed half the risk of AD that nonsmokers had. The odds ratios seen here are relatively consistent; they reflect a decreased risk for AD in smokers (Lee 1994). Many of the 95% CIs include 1.0, an indication that “no association” cannot be excluded as an explanation. Wide CIs can be caused by small sample sizes. However, combining studies to increase overall sample size increases the power to detect an association when it exists; it also leads to narrower CIs. Because of the consistent odds ratios and evaluation of other evidence available, Lee was justified in conducting a meta-analysis of the data. The net effect observed by meta-analysis was a significant 40% reduction in risk of AD among smokers (Lee 1994). Taken as the sum of many studies, this
evidence seemed convincing. Table 1 shows the principal case-control studies reviewed by Lee and the estimated relative risk for AD associated with smoking.

Summary measures of effect based on a meta-analysis alone may be misleading. The biases inherent in each of the individual studies can influence the summary estimate. Therefore, though it takes advantage of the large numbers of patients to increase power and narrow CIs, combined analysis does not adjust or eliminate inaccurate results (i.e., due to bias).

Consider the case-control design as shown in Figure 1. The investigator selects cases (persons with AD) and control subjects (persons without AD) and then attempts to determine the smoking history before disease onset in cases and, at a similar time, for the persons called control subjects. If the way in which cases or control subjects are identified and enrolled is related to smoking history, then some bias is likely to exist. If exposure information is obtained differently for cases relative to control subjects, then another bias is likely to exist. Suppose that an investigator carefully included cases who were considered as exposed only if they had a smoking history of at least 2 years before any AD symptoms but, for control subjects, captured their current smoking habits rather than at a comparable point in the past. This could systematically include more control subjects relative to smokers and therefore bias the association.

Recall bias is a well-known contaminant of case-control studies. It stems from differential recall or reporting of exposure status by cases and control subjects. For AD this is a well-recognized problem because AD patients all have memory deficits. Reports given by proxy (spouse, child, other) are often used to obtain smoking history. However, if proxy reports are not used for the control patient as well, then bias could result. Objective exposure measurement (e.g., through specific records or biological measurements) is ideal but difficult to attain.

A more subtle form of bias can result if a cross-sectional sample of AD cases is selected for inclusion in the study. For example, at the start of the case-control study the investigator may attempt to enroll all the living and available AD patients seen in an Alzheimer’s clinic during the last 5 years. This is often described as enrolling prevalent rather than “incident” or new cases. When survival among persons with the disease is related to the exposure of interest, then bias can result. Figure 2 shows a hypothetical representation of the life span of AD patients following diagnosis. Each line represents an individual case from diagnosis to death. The short lines represent short time spans and the long lines represent a greater time between diagnosis and death. Notice that the vertical line representing the cross-sectional sample of existing cases intersects with more longer duration cases than shorter duration ones. Persons who survive only a short time with the disease may differ in many ways from those who survive longer. For example, they may have greater disease severity or more comorbidity. Assessing risk factors primarily from longer surviving cases would

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Odds ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Katzman et al 1989</td>
<td>1989</td>
<td>0.2</td>
<td>0.1–0.5</td>
</tr>
<tr>
<td>Grossberg et al 1989</td>
<td>1989</td>
<td>0.3</td>
<td>0.2–0.6</td>
</tr>
<tr>
<td>Ferini-Strambi et al 1990</td>
<td>1990</td>
<td>0.4</td>
<td>0.2–0.8</td>
</tr>
<tr>
<td>Brenner et al 1993</td>
<td>1993</td>
<td>0.6</td>
<td>0.4–1.0</td>
</tr>
<tr>
<td>Chandra et al 1987</td>
<td>1987</td>
<td>0.6</td>
<td>0.3–1.4</td>
</tr>
<tr>
<td>French et al 1985</td>
<td>1985</td>
<td>0.7</td>
<td>0.3–1.5</td>
</tr>
<tr>
<td>Amaducci et al 1986</td>
<td>1986</td>
<td>0.7</td>
<td>0.4–1.2</td>
</tr>
<tr>
<td>Graves et al 1990, 1991</td>
<td>1990</td>
<td>0.8</td>
<td>0.5–1.2</td>
</tr>
<tr>
<td>Dewey et al 1988</td>
<td>1988</td>
<td>0.8</td>
<td>0.4–1.6</td>
</tr>
<tr>
<td>Broe et al 1990</td>
<td>1990</td>
<td>0.8</td>
<td>0.5–1.3</td>
</tr>
<tr>
<td>Heyman et al 1984</td>
<td>1984</td>
<td>1.0</td>
<td>0.5–2.0</td>
</tr>
<tr>
<td>Barclay and Kheyfets 1989</td>
<td>1989</td>
<td>1.3</td>
<td>0.5–3.2</td>
</tr>
<tr>
<td>Shalat et al 1987</td>
<td>1987</td>
<td>1.3</td>
<td>0.7–2.3</td>
</tr>
<tr>
<td><strong>Summary estimate (Lee 1994)</strong></td>
<td></td>
<td><strong>0.6</strong></td>
<td><strong>0.5–0.8</strong></td>
</tr>
</tbody>
</table>

CI, confidence interval.

Figure 1. Simple representation of the case-control study design. Hx, history; AD, Alzheimer’s disease.

Figure 2. Cross-sectional sampling of prevalent cases preferentially includes patients with longer survival after diagnosis (also known as prevalent case bias or length-biased sampling).
not allow the investigator to distinguish between factors related to disease onset and factors related to survival.

One way to evaluate whether biases may have been present in the case-control studies already discussed is to compare their results to those obtained through a study design that is less susceptible to those biases. (Keep in mind that small $p$ values/significance levels are only a probabilistic statement of chance and can result in either biased or unbiased results.) Figure 3 shows a simplified diagram of a cohort study design. Persons without AD are first classified in relation to their exposure to smoking and then reassessed for development of AD, at specified intervals thereafter. This allows observation of incident cases and avoids the prevalent case bias. Because exposure is assessed before onset of disease there is also less chance of reference year bias or recall bias. A major threat to the validity of cohort studies is loss of subjects during the follow-up period. The results of several current cohort studies are shown in Table 2.

Wang et al (1999) conducted an important and informative population-based study in 1987 in Stockholm: the Kungsholmen project, in which persons over age 74 were enrolled. A case-control and a cohort analysis were conducted. Screening of the initial patients revealed 198 prevalent dementia cases, which were eliminated from the incidence portion of the study. These became subjects in a case-control analyses. After exclusion of patients with cognitive impairment and dementia, 343 cognitively intact patients entered the cohort study. From the case-control analyses using the 198 prevalent cases, as compared with nondemented patients of similar age, Wang et al (1999) found that smoking appeared to be a protective factor for AD (odds ratio = 0.6; 95% CI, 0.4–1.1). This result is quite similar to the case-control studies cited previously and to Lee’s summary table (Table 1). However, results obtained from the cohort study showed no decreased risk of AD among smokers (hazard ratio = 1.1; 95% CI, 0.5–2.4). Based on the cohort, Wang then studied mortality among the demented and nondemented, and compared smokers with nonsmokers within each group. It was found that, among the demented, smokers were 3.4 times more likely to die during a 5-year follow-up than demented nonsmokers. In contrast, among the nondemented, smokers’ risk of death differed much less from that of nonsmokers of similar age.

Figure 2 showed the potential that longer duration cases tend to be most included in case-control studies. Wang’s results establish that such a bias may be present in at least some of the case-control studies. Wang et al (1999) concluded that persons with AD who are also smokers often die more quickly and therefore would be unavailable for case-control studies. This would give the appearance that smoking among cases was much less common than it actually is and, in turn, that smoking among control subjects was more common than among cases. Thus, a case-control study could conclude that smoking has a “protective” effect.

Ott et al (1998) reported on another cohort study that was conducted in Rotterdam and included all eligible persons over 55 years of age. Six thousand eight hundred seventy persons who provided information concerning smoking were observed for incidence of dementia and AD. The length of follow-up was approximately 2 years at this report. In comparing former smokers and “never” smokers, Ott et al (1998) observed a relative risk of 1.4 (95% CI, 0.9–2.0). Current smokers enjoyed a relative risk of 2.2 (95% CI, 1.3–3.6) relative to never smokers. This study indicates a twofold increased risk for AD among smokers; men showed higher risk than women. Ott et al then examined the simultaneous effects of smoking and apolipoprotein E (APOE) genotype. Among persons with an $e4$ allele (a risk factor for AD), in either current or former smokers, the effect of smoking was null. Among persons without an $e4$ allele the risk of AD due to smoking appeared elevated. The biological reasons for effect modification (interaction) of the smoking–AD association by APOE $e4$ are not obvious. Apolipoprotein E is involved in

---

**Figure 3.** Simplified representation of a cohort study design to assess the association between smoking and the onset of Alzheimer’s disease (AD).

---

**Table 2.** Selected Cohort Studies of the Association between Smoking and Alzheimer’s Disease

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Relative risk</th>
<th>CI</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doll et al 2000</td>
<td>2000</td>
<td>0.78 (former)</td>
<td></td>
<td>0.6–1.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.83 (current)</td>
<td></td>
<td>0.6–1.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.99 (current)</td>
<td></td>
<td>0.8–1.3</td>
</tr>
<tr>
<td>Launer et al 1999</td>
<td>1999</td>
<td>1.2 (former)</td>
<td></td>
<td>0.8–1.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.7 (continuing)</td>
<td></td>
<td>1.2–2.5</td>
</tr>
<tr>
<td>Merchant et al 1999</td>
<td>1999</td>
<td>0.7 (former)</td>
<td></td>
<td>0.5–1.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.7 (continuing)</td>
<td></td>
<td>1.1–2.8</td>
</tr>
<tr>
<td>Ott et al 1998</td>
<td>1998</td>
<td>1.4 (former)</td>
<td></td>
<td>0.9–2.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.2 (continuing)</td>
<td></td>
<td>1.3–3.6</td>
</tr>
<tr>
<td>Wang et al 1999</td>
<td>1999</td>
<td>1.1</td>
<td></td>
<td>0.5–2.4</td>
</tr>
</tbody>
</table>

CI, confidence interval.
cholesterol transport, but exactly how APOE acts in the pathogenesis of AD is not known. One might speculate that both APOE4 and smoking act on cholesterol level, a risk factor for vascular disease, but in the presence of APOE4 the additional effect of smoking on cholesterol level is less. Then, one must also hypothesize a vascular component to AD to complete the mechanism. Much more foundation is needed to establish the true nature of these observed associations.

Merchant et al (1999) conducted a smaller cohort study in New York City; it included 1062 persons, who were observed for approximately 2 years. Follow-up examination and diagnosis were conducted in a manner similar to that of Ott et al (1998). Table 2 shows the risk of AD in former smokers relative to “never” smokers was consistent with a conclusion of “no association.” However, when “current” smokers (smoking at baseline) were compared with never smokers, a modest increased risk was seen for current smokers. The relative risk estimates presented were adjusted for education and ethnicity. Merchant et al (1999) was also able to stratify by APOE €4–containing genotype. The result was consistent with that obtained by Ott et al (1998): increased risk was seen among those with non-€4-containing genotypes, but essentially a null association between smoking and AD among those who carry the €4 allele.

Launer et al (1999) reported an analysis of data compiled from similar cohort studies conducted in Denmark, France, The Netherlands, and the United Kingdom—the Eurodem cohort. The contributing cohorts were designed and conducted under the principle that their data would be combined, and thus noncomparability across study sites was not a major detriment (although adjustment for study site was included in analyses). The analysis included 13,147 subjects, who were observed for approximately 2 years, resulting in 24,731 person-years of observation. As with the Ott et al (1998) and Merchant et al (1999) studies, no increased risk of AD due to smoking was seen in former smokers relative to “never” smokers. Modestly increased risk was observed, however, for those who were smokers at baseline when compared with never smokers. Concerned that the effect of increased risk could be due to vascular disease, Launer also analyzed the data excluding those AD patients with contributing cardiovascular disease; the results did not change materially. Analysis by gender showed somewhat greater relative risk estimates among men than among women. However, there was not a sufficient increase to conclude that an interaction between gender and smoking (effect modification by gender) was a credible conclusion. When the data were stratified by “family history” of dementia, the increased risk in current smokers was confined to the groups designated as having no family history. These results are analogous to the findings of APOE genotype (Merchant et al 1999; Ott et al 1998) but may represent a different phenomenon because family history may be a surrogate for many genetic and environmental factors.

Doll et al (2000) reported the association between smoking and AD, but based diagnosis of AD on whether dementia was listed as a cause of death on the death certificate. This study was based on the original study of British physicians that began in 1951 and included 34,439 male patients at the outset. Follow-up was performed at 6- to 12-year intervals following study initiation. Patients’ current smoking status was recorded as smoking for 10 years or more before death, and former smokers had stopped smoking roughly 34 years before death. This study relied on death certificates for the diagnosis of dementia and AD. Doll et al were quite aware that death certificates could be inadequate to determine the type of dementia. Therefore, they conducted a substudy of the death certificates from 1996 to 1998. In the substudy, additional information was obtained from certifying physicians or other hospital physicians concerning dementia diagnosis. Results of the substudy were then applied to estimate the proportion of cases with death certificate entries indicating dementia that were likely to be AD. As of December 1998, 24,133 deaths had occurred from the original cohort; this constituted the base of the Doll study.Death certificates identified only 473 subjects as demented. The substudy results allowed those 473 to be further classified as AD or as “other” dementias. As shown in Table 2, there was no increased risk of AD associated with smoking (based on death certificate diagnosis of AD).

Although this study has many distinct advantages, such as long-term follow-up and careful exposure measurement well before disease onset, it also causes some concern about diagnostic misclassification. The substudy of dementia diagnosis based on death certificate entries was an excellent step toward validity. Doll does not propose a method to detect unrecognized dementia and AD prevalence among the approximately 24,000 other deaths that had occurred at time of analysis. The reliance on death certificate entries alone to identify cases may have failed to capture many patients with AD. For example, persons who die of cancer while in the early to moderate stages of dementia may not have dementia entered as a cause of death on their death certificate.

Calendar year of death may have also influenced the likelihood that AD or dementia would be recorded on a death certificate (Doll et al 2000). Death certificates completed between 1951 and 1978 may be considerably less likely to include dementia as a cause, since diagnostic fashion tended to regard some “senility” as a part of aging. This study contrasts others cited previously, in that its outcome is not incidence of the clinical diagnosis of
dementia and AD. Nondifferential misclassification of diagnosis (missing persons who died with unrecorded dementia, irrespective of smoking) could be expected to drive an observed association toward the null.

Discussion

We have discussed some of the potential biases associated with case-control studies, and all the results appeared rather consistent. History of smoking was more frequent among control subjects than it was among cases. Stated causally, smoking appeared to be protective for AD when evidence was limited to case-control studies. Biases inherent in selection of patients, such that exposure to smoking was favored in control subjects, through differential survival (Wang et al 1999), appear to be one reasonable and credible explanation for the results of the case-control studies. Bias in the exposure information from cases and control subjects could also account for some of the observed effect. Although confounding effects of variables can be treated with multivariate statistical methods in the analysis, bias is essentially immune to statistical therapy.

Cohort studies offer advantages to the study of smoking and AD. The exposure information can be determined before onset of disease, thus avoiding recall bias. Incidence dementia and AD can be observed directly if there is sufficient, detailed follow-up examination sensitive to the detection of dementia. The interval between follow-up visits, if relatively long (e.g., 5–10 years), may cause new cases of disease to be missed. A supplementary surveillance system can sometimes be implemented to avoid loss of diagnostic information.

Relatively short follow-up time is a characteristic of most of the cohort studies presented. As Doll et al (2000) point out, long and extended follow-up is preferable because it may give a more complete picture of the lifelong influence of exposure on disease (not, however, a long interval between follow-up exams). With AD, the optimum exposure–disease interval is unknown. This means that we do not know which exposures occurred, but which were nonetheless irrelevant, because they could not result in disease onset. We are also not able to reliably estimate which exposures may be irrelevant to etiology, because they have occurred following the biological onset of disease.

A 2-year follow-up seems short, and we would be unsure about whether the biological disease process had already begun when the patients were enrolled and, therefore, whether the exposure being measured occurred sufficiently beforehand. There is also the concern that the first follow-up period may contain more patients who had biological disease at entry. Successive follow-up examinations over time would lead to a more stable incidence of disease and possibly more valid exposure–disease associations.

The length of follow-up in a cohort study is viewed somewhat differently than follow-up in an RCT treatment study would be. The time axis for the cohort study of dementia is best viewed as age rather than accrued time since entry, because age dramatically increases the risk of AD. An RCT, on the other hand, provides the “treatment” to patients at the beginning of the study and then observes its effect; thus, length of follow-up relative to treatment initiation may be critical to observe an effect. In a cohort study, however, patients experience exposure to smoking and age for various lengths of time before observation begins. Thus, as long as the power of the study is sufficient and the effect of biases in diagnosis or differential loss to follow-up is minimal, the observed associations should be similar to those obtained after longer follow-up.

Death certificate diagnosis (e.g., Doll et al 2000) of dementia causes many concerns regarding the extent of “missed” disease cases and the potential exposure association within those missed. Consider that in the Eurodem cohort study 13,147 persons were observed for about 2 years and 277 incident AD cases were seen (Launer et al 1999). In contrast, the study of male British doctors observed 34,439 individuals from 1951 to 1998 and saw 473 cases of dementia diagnosed by death certificate entry (including 370 AD), out of 24,133 total deaths (Doll et al 2000). Thus, observed dementia incidence in the Doll study appears to be abnormally lower than would be expected from the Eurodem data.

The methodological superiority of cohort studies to eliminate well-known biases (prevalence and recall) causes reinterpretation of earlier case-control study findings. “Consistent” case-control study findings are now viewed as potentially systematically biased. Findings from the cohort studies discussed are also consistent but contrast with the case-control results. Loss to follow-up is usually the greatest threat to the validity of a cohort study, but in these studies low loss to follow-up does not appear to have compromised validity. How exposure to smoking is classified (as current or former) may have some importance in determining whether increased risk might exist. Certainly if an etiologic time window could be described for the exposure–disease relationship, investigators would be able to more accurately determine whether the effect of smoking on AD was positive, negative, or null. In sum, the cohort studies to date lean toward some increased risk for those continuing to smoke versus those who never smoked, and no association for those who quit in the past. Doll et al (2000) point out that it may be more reasonable to group exsmokers who stopped a long time ago with nonsmokers to form the reference group. In the Doll study, exsmokers quit about 34 years before death, and therefore,
smoking so long in the past may be unrelated to disease etiology. Unfortunately, determination of the temporal relation, which renders some exposures relevant to disease onset and others irrelevant, has not yet been reached.

Although these cohort studies show smoking may be unrelated to the onset of AD or could increase risk of AD, they do not specifically address the potential long-term effects that might be obtained from the modulation of nicotinic receptors in AD (Newhouse et al 2000). Cigarette smoke contains many chemicals and substances, making it impossible to evaluate how nicotinic modulation alone could alter AD risk.

Aspects of this work were presented at the symposium “Nicotine Mechanisms in Alzheimer’s Disease,” March 16–18, 2000, Fajardo, Puerto Rico. The conference was sponsored by the Society of Biological Psychiatry through an unrestricted educational grant provided by Janssen Pharmaceutica LP.

The author’s work is also supported in part by grants from the National Institute on Aging (AG06781, AG07584, AG16976).

References


