Dose-Response: An International Journal

Volume 10
Issue 2 *Special Issue on the Role of Linear and Nonlinera Dose-Response Models in Public Decision-Making*

Article 9

6-2012

LOW-DOSE NONLINEAR EFFECTS OF SMOKING ON CORONARY HEART DISEASE RISK

Louis Anthony (Tony) Cox, Jr
Cox Associates

Dose-Response, 10:219–232, 2012
Formerly Nonlinearity in Biology, Toxicology, and Medicine

ISSN: 1559-3258
DOI: 10.2203/dose-response.11-038.Cox

International Dose-Response Society
www.Dose-Response.org

LOW-DOSE NONLINEAR EFFECTS OF SMOKING ON CORONARY HEART DISEASE RISK

Louis Anthony (Tony) Cox, Jr. □ Cox Associates

□ Some recent discussions of adverse human health effects of active and passive smoking have suggested that low levels of exposure are disproportionately dangerous, so that "The effects of even brief (minutes to hours) passive smoking are often nearly as large (averaging 80% to 90%) as chronic active smoking" (Barnoya and Glantz, 2005). Recent epidemiological evidence (Teo *et al.*, 2006) suggests a more linear relation. This paper reexamines the empirical relation between self-reported low levels of active smoking and risk of coronary heart disease (CHD) in public-domain data from the National Health and Nutrition Examination Survey (NHANES). Consistent with biological evidence on J-shaped and U-shaped relations between smoking-associated risk factors and CHD risks, we find that low levels of active smoking do not appear to be associated with increased CHD risk. Several methodological challenges in epidemiology may explain how model-derived estimates can predict low-dose linear or concave dose-response estimates, even if the empirical (i.e., data-based) relation does not show a clear increased risk at the lowest doses.

Keywords: Coronary Heart Disease (CHD), hormesis, U-shaped, J-shaped, empirical dose-response model, confounding, modeling bias, classification tree analysis

INTRODUCTION: DOES HORMESIS FAIL FOR SMOKING AND CORONARY HEART DISEASE?

An emerging working hypothesis for some toxicologists and risk assessors is that many – perhaps most – biological dose-response relations exhibit J-shaped or U-shaped regions at low doses. That is, probability of harm (or, more generally, of exposure-related departures of variables from their "normal" levels) decreases with increasing dose at sufficiently small exposure levels, even if it increases with increasing doses at higher exposure levels. When this pattern holds, responses to low levels of exposures cannot necessarily be extrapolated from observed dose-response relations at higher doses.

Although considerable empirical support has been advanced in support of this "hormesis" hypothesis (Calabrese and Baldwin, 2001), the universality of its application is still being assessed. The shape of dose-response functions for complex mixtures, such as diesel exhaust or cigarette smoke, can potentially be especially valuable in either supporting the hormesis hypothesis or in understanding how it breaks down.

Address correspondence to Dr. Tony Cox, Cox Associates, 503 Franklin Street, Denver, Colorado, 80218; Email: tcoxdenver@aol.com

In contrast to hormesis, studies of environmental tobacco smoke (ETS) suggest that its effects on risks of diseases such as CHD are much *larger* than would be expected based on associations in active smokers exposed to much higher doses. Although some commentators have construed this mismatch as suggesting that perhaps reported ETS-CHD associations reflect incompletely controlled confounders or statistical modeling biases (e.g., Nilsson, 2001), others suggest that active smokers might have adaptive responses that create less-than-proportional increases in CHD risk at relatively high exposure levels compared to the increases in risks experienced by non-smokers from relatively low ETS exposures (Glantz and Parmley, 1991; Wells, 1994; Law *et al.*, 1997; Law and Wald, 2003). For example, it has been stated that "Evidence is rapidly accumulating that the cardiovascular system – platelet and endothelial function, arterial stiffness, atherosclerosis, oxidative stress, inflammation, heart rate variability, energy metabolism, and increased infarct size – is exquisitely sensitive to the toxins in secondhand smoke. The effects of even brief (minutes to hours) passive smoking are often nearly as large (averaging 80% to 90%) as chronic active smoking" (Barnoya and Glantz, 2005). If true, this provides an important counter-example to the hormesis hypothesis for smoking and CHD risk.

This paper reexamines the empirical relation between relatively low levels of active smoking and CHD risks. It seeks to reassess whether the hypothesis of hormesis must be rejected in this context. A key methodological challenge is that selecting particular statistical and epidemiological modeling assumptions can strongly affect the results obtained and the resulting interpretation of available low-dose smoking-CHD data, suggesting a need for multibias modeling (Greenland, 2005). To address this challenge, we consider non-parametric techniques for identifying possibly nonlinear dose-response relations in large epidemiological data sets.

The following sections briefly review key results from the relevant epidemiological and biological literatures and then present a new analysis of a publicly available data set, the National Health and Nutrition Examination Survey (NHANES) study, that provides data on active smoking and CHD risks, among other outcomes.

PREVIOUS FINDINGS AND NEED FOR IMPROVED METHODS TO QUANTIFY LOW-DOSE EFFECTS OF SMOKING ON CHD RISK

Since the 1960s, the relation between risk of coronary heart disease and exposures to relatively low levels of active cigarette smoking (e.g., five or fewer cigarettes per day) or environmental tobacco smoke (ETS) has been examined in dozens of studies and hundreds of publications, resulting in diverse epidemiological and biomedical findings, data interpretations, meta-analyses, and reviews. Table 1 summarizes examples of various conclusions from studies on ETS and CHD risk. Some studies

Nonlinear smoking effects on CHD risk

TABLE 1. Examples of Reported Findings on ETS and CHD Risk

Study	Findings*
Chen *et al.*, 2004	"When all CHD categories are combined there is a regular, significant gradient [dose-response for ETS and CHD]... [But] there was a higher prevalence of questionnaire angina, undiagnosed CHD, and all CHD in subjects with no detectable cotanine..."
Enstrom and Kabat, 2003	"For participants followed from 1960 until 1998 the age adjusted relative risk (95% confidence interval) for never smokers married to ever smokers compared with never smokers married to never smokers was 0.94 (0.85 to 1.05) for coronary heart disease...No significant associations were found for current or former exposure to environmental tobacco smoke before or after adjusting for seven confounders."
Nilsson, 2001	"**By pooling data** from 20 published studies on ETS and heart disease, some of which reported higher risks than is known to be caused by active smoking, **a statistically significant association with spousal smoking is obtained.** However, in most of these studies, many of the **most common confounding risk factors were ignored and there appears to be insufficient evidence to support an association between exposure to ETS and CHD.**"
Law *et al.*, 1997	"Cohort and case control studies show a **30% excess risk of ischemic heart disease in nonsmokers whose spouses smoke** compared with that in nonsmokers whose spouses do not smoke. **There is a nonlinear dose-response**; the excess risk from actively smoking 20 cigarettes/day is only 80%. ...In experimental studies passive and active smoking have similar effects on platelet aggregation. The collective evidence supports a significant effect of low dose tobacco smoke exposure in causing ischaemic heart disease."
Steenland *et al.*, 1996	"Results are consistent with prior reports that never-smokers currently exposed to ETS have about 20% higher CHD death rates. However, **our data do not show consistent dose-response trends** and are possibly subject to confounding by unmeasured risk factors."
Gori, 1995	"Numerous epidemiologic studies report that **the active smoking of less than 10 cigarettes/day is not associated with measurable risk of coronary heart disease (CHD)**. Thus, even assuming that ETS and MS [mainstream] have equivalent biologic activities, conceivable **ETS doses to nonsmokers are far below apparent no-effect thresholds for active smoking**."
Wells, 1994	"The effects of passive smoking on ischemic heart disease are reviewed. Short-term exposures of 20 min to 8 h result in increased platelet sensitivity and decreased ability of the heart to receive and process oxygen. Longer term exposure results in plaque buildup and adverse effects on blood cholesterol. The available epidemiology is reviewed, and it is concluded that **passive smoking increases the coronary death rate among U.S. never smokers by 20% to 70%.**"
Glantz and Parmley, 1991	"**Nonsmokers appear to be more sensitive to ETS than do smokers,** perhaps because some of the affected physiological systems are sensitive to low doses of the compounds in ETS, then saturate, and also perhaps because of physiological adaptions smokers undergo ... These results **suggest that heart disease is an important consequence of exposure to ETS**."

*All emphases added.

have reported elevated risks even at the lowest doses studied (e.g., Njolstad *et al.*, 1996 for 1-9 cigarettes per day in a Norwegian population), while others have reported either no detected independent effect of smoking on CHD risks (Chien *et al.*, 2005 for a community in Taiwan) or elevated risks only for smoking exposures above a threshold, such as 20 pack-years (Lee *et al.*, 2001, for Chinese, Malay, and Asian Indian males in Singapore).

Evidence of Hormesis for Smoking and CHD Risk Factors

Several biological mechanism have been proposed for how low-level or second hand smoke exposure can increase CHD risk. However, careful reviews of the relevant biology (e.g., MacCallum, 2005) indicate that many of the biological effects of smoking are associated with – but do not necessarily cause – increased CHD risk. These effects may serve as biomarkers of exposure and of CHD risk, but without necessarily constituting causal mechanisms implying that exposure increases risk. For example, "Despite their evident importance in MI [myocardial infarction], it has proven difficult epidemiologically to demonstrate associations between platelet properties and CHD events in prospective studies", perhaps because of measurement difficulties and uncertainties (MacCallum, 2005, p. 37). Similarly, is not clear whether much-discussed markers such as C-reactive protein (CRP) actually *cause* any increases in CHD risk, although CRP is *predictive* of CHD risk – perhaps because it is caused by other conditions that also cause increased CHD risk (e.g., Miller *et al.*, 2005).

Many important biological variables that are related to smoking and/or CHD risk have been found to exhibit hormetic (U-shaped or J-shaped) relations with CHD risk (and, in many cases, with all-cause mortality risk). Examples include: C-reactive protein (O'Callaghan *et al.* 2005), total and low-density lipoprotein cholesterol in a prospective study among elderly men (Curb *et al.*, 2004); serum insulin levels in a cross-sectional survey of 500 men and 500 women aged 40-79 years in Italy (Bonora *et al.*, 1998); blood pressure in the first two years following MI in men 45-57 years old (Flack *et al.*, 1995); hematocrit among women in the Framingham heart study (Gagnon *et al.*, 1994); and possibly heart rate for sudden CHD death, although the evidence for this is mixed (Dyer *et al.*, 1980). Thus, even bearing in mind that subgroup analyses can create false positives (Brookes *et al.*, 2001), it appears that current biological knowledge of CHD etiology allows the possibility of a U-shaped or J-shaped relation between exposures (including smoking) that affect one or more of these variables (Hatsukami *et al.*, 2005) and resulting risks of CHD. Therefore, empirical assessment of the true shape of the low-dose dose-response relation based on epidemiological data remains an important and worthwhile challenge.

EMPIRICAL ASSESSMENT OF THE LOW-DOSE RELATION BETWEEN SMOKING AND CHD RISK IN THE NHANES STUDY

To investigate the empirical (i.e., data-driven) shape of the relation between relatively low exposures to cigarette smoking and resulting risk of CHD without making any strong *a priori* parametric modeling assumptions, we downloaded survey data from the National Health and Nutrition Examination Survey (NHANES) study for 2001-2002 (on-line at www.cdc.gov/nchs/about/major/nhanes/nhanes01-02.htm#Examination%20Files). As emphasized in the Analytic Guidelines for this survey, NHANES is a complex survey sample, and careful attention to weighting is needed to understand, interpret, and generalize from it to other (e.g., national) populations. However, in the analyses in this section, we only use conditional (internal to the survey data set) relations between smoking levels and CHD risks. These conditional relations can be obtained directly from the survey data, and no attempt to generalize to other populations is made here.

Figure 1A plots the mean proportions (and 95% binomial confidence intervals) of subjects who answered "Yes" to the question "Has a doctor or

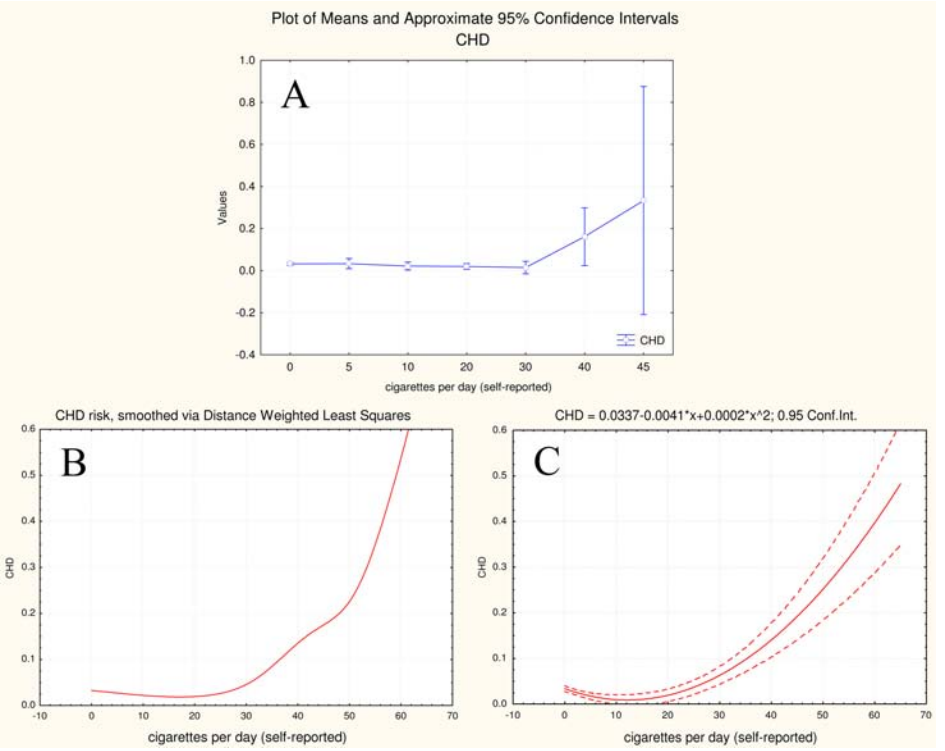

FIGURE 1. A: CHD Risk *vs.* Self-Reported Smoking Levels in the NHANES Survey. Each positive number on the x axis represents the range of values that are closest to it. "45" is a code for "more than 40". B: Nonparametric Regression of CHD Risk vs. Cigarettes/Day. C: Polynomial Regression of CHD Risk vs. Cigarettes/Day. Dashed lines indicate approximate lower and upper 95% confidence limits.

other health professional ever told you [or subject] that {you/s/he} had coronary heart disease?" as the dependent variable (vertical axis). Self-reported smoking level is the explanatory variable (horizontal axis). The data in the plot suggest no clear increase in CHD risk at relatively low levels of smoking exposure. (Breaking down the subjects by sex and by age groups, such as 65 or older *vs.* younger, does not change these conclusions.) Figure 1B shows a nonparametric regression curve (fit by distance-weighted least squares) used to smooth the scatter plot of CHD risk indicator values (0 = no, 1 = yes) *vs.* cigarettes per day. Figure 1C fits a polynomial (quadratic) regression curve to the same data. In this parametric model, the nonlinear (J-shaped) quadratic term is statistically significant. The estimated minimum risk occurs at about 11 cigarettes per day, consistent with earlier findings in multiple data sets (reviewed in Gori, 1995) of no apparent significant increase in CHD risk among smokers of 10 cigarettes per day or less.

Definition of the "zero" exposure level plays a potentially important role in such statistical modeling of low-dose effects. In these figures, the "0" level was reserved for subjects who reported not having smoked more than 100 cigarettes throughout their lives (and who report being non-smokers now.) This exposure category may be associated with other behaviors (e.g., healthy diet, exercise, etc.) that confound the effects of the zero smoking level. Nonetheless, Figure 1 suggests that, while 40 or more cigarettes per day is clearly associated with a significantly increased CHD risk, there is no such clear, significant increase at the lowest reported levels.

Figure 2 repeats the preceding analysis for ex-smokers who report having smoked at least 1 cigarette per day when they used to smoke. Again, the main conclusion from these plots is that, for relatively low level of past smoking (e.g., below 10 cigarettes per day), higher reported numbers of cigarettes per day do not appear to be associated with increased CHD risk; to the contrary, a U-shaped pattern appears to be possible.

The nonlinear dose-response relations suggested in these figures may not reflect a true cause-and-effect relation, but rather the effects of confounders (such as alcohol and coffee consumption, which, as noted previously, are associated with cigarette consumption and also have U-shaped relations with CHD risk) or other variables associated with both smoking and CHD risks.

DISCUSSION

The following paragraphs summarize technical issues that can drive differences in conclusions between epidemiological studies such as those in Table 1. These threats to valid inference must be overcome to create defensible estimates of dose-response relations from epidemiological data.

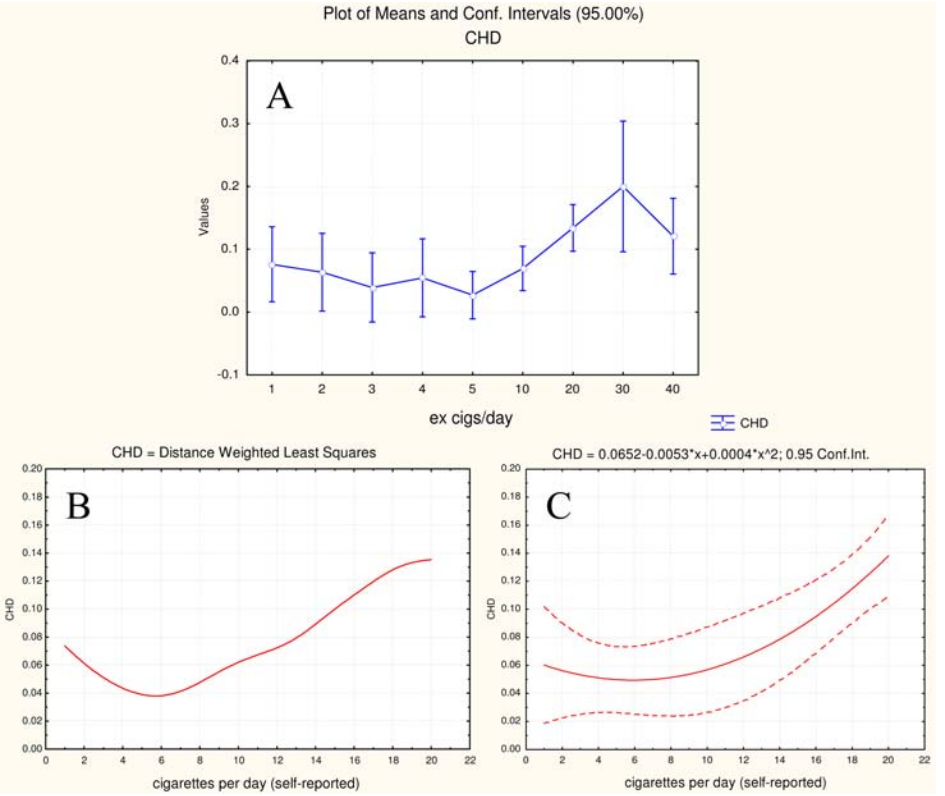

FIGURE 2. A: CHD vs. Past Smoking of 1-20 cigarettes per day. B: Nonparametric Regression, CHD vs. Past Smoking of 1-20 cigarettes per day. C: Polynomial Regression CHD vs. Past 1-20 cigarettes per day.

Model Form Selection

The choice of a statistical model constrains what it can reveal about low-dose effects of exposures. For example, many investigators have relied on specific parametric or semi-parametric models (e.g., logistic regression or Cox proportional hazards models) to interpret epidemiological data on smoking exposures and CHD risks. Such models imply that model-estimated risks must be elevated at low doses if they are elevated at high doses, whether or not this is what the data show. Some epidemiological studies that have reported elevated risks at relatively low exposure levels (e.g., Njolstad *et al.*, 1996) have done so based on such statistical model implications.

Older analyses that report raw data rather than such statistical model-based estimates can show patterns for CHD risks at low exposure levels strikingly different from the elevated risks predicted by statistical models that assume a single set of coefficients for all exposure levels. For example, Freund *et al.* (1993)'s analysis of Framingham Heart Study data (their Table 2) shows CHD rates that are *lower* among men who smoke 1-10 cigarettes per

day than among never-smokers, for both age groups considered. Similarly, Bush and Comstock (1983) show reduced risks of CHD in the lowest exposure group considered (smokers of 1-9 cigarettes per day) in each of three age groups examined (25-44 years, 45-64 years, and 65-74 years) in a study of smoking and CHD mortality risks in women, even after adjusting for variables (e.g., marital status, education, housing quality, and frequency of church attendance) that have been associated with CHD risk in other data sets. Jenkins *et al.,* (1968, Table 1 and 2) show a slightly smaller rate of CHD cases among smokers of 1-15 cigarettes per day than among non-smokers (12.2 vs. 13.3 per 1000 men per year) for men aged 50-59 years at study intake, but a greater rate of CHD cases among smokers of 1-15 cigarettes per day than among non-smokers (5.3 vs. 3.7 per 1000 men per year) for men aged 39 to 49 at study intake. The reported results do not have enough resolution to examine rates specifically for 1-10 cigarettes per day to allow comparison with the results of Freund *et al.* (1993), but it seems plausible that CHD risks could be elevated among smokers of 10 or more cigarettes per day, even if not among smokers of fewer than 10 cigarettes per day (the apparent threshold for detectably elevated risk mentioned by Gori, 1995.)

Variable Selection

An enduring challenge in multivariate risk modeling is selection of variables to include in the final model. Different logistic regression models that appear to make equally statistically valid selections of predictor variables can give very inconsistent predictions for outcomes such as mortality following MI (Steyerberg *et al.*, 2004). Automated variable-selection techniques can result in models that are unstable and not reproducible (e.g., across bootstrap samples of the original data set) (Austin and Tu, 2004), while manual selection of variables by investigators, especially with preconceived theories, may lead to biased conclusions and to errors that are not readily apparent based on standard statistical tests (e.g., Greenland, 2005).

Empirically, a study that applied several variable-selection algorithms (including logistic regression with forward or backward stepwise variable selection; neural networks; self-organizing maps; and rough sets) to the same data set of 500 records with 45 predictor variables from patients with chest pain, and a dichotomous dependent variable indicating whether myocardial infarction (MI) occurred, showed that the different variable-selection methods yielded inconsistent results. Only one variable out of 45 (ST elevation) was selected by all methods. Only two out of eleven variable-selection methods (one of which was expert cardiologist opinion) identified smoking as a useful predictor of MI risk for patients in this data set (Dreiseitl *et al.*, 1999). Such inconsistencies may help to explain differences in conclusions and reported associations among studies that use different methods of variable selection.

Variable Coding

How continuous variables are coded into discrete ranges or levels can also affect statistical conclusions about associations. For example, as mentioned above, Jenkins *et al.* (1968) found evidence of a dose-response relation for CHD risk that progressively increases with increasing dose in the younger of two age-groups analyzed, taking 1-15 cigarettes/day as the lowest range considered; but Bush and Comstock (1983) showed reduced risks of CHD in all age groups analyzed, taking 1-9 cigarettes/day as the lowest range considered. The reported findings may be sensitive to the modelers' choices of how to bin the exposure and covariate data. (For this reason, it is often recommended in modern statistical and epidemiological methodology that continuous variables should not be artificially coded into discrete levels; see e.g., Stromberg, 1996; Royston *et al.*, 2006).

Confounding

The etiology of CHD is complex and still imperfectly elucidated. A surprising variety of risk factors can potentially act as confounders, associated both with smoking and with increased CHD risk. Studies in multiple countries have shown that many risk factors (e.g., low exercise, poor diet, high body fat, high blood pressure, low concentration of high density cholesterol, low income, low education, low cognitive performance) tend to cluster with each other and with both smoking and CHD risk in the same individuals, often starting relatively early in life (e.g., by adolescence) and persisting thereafter (e.g., Ebrahim *et al.*, 2004.) Smoking is positively associated with coffee and alcohol consumption, both of which have been reported to have hormetic (J-shaped or U-shaped) dose-response relations with CHD risk (Kleemola *et al.*, 2000; Murray *et al.*, 2002).

Various psychosocial factors are also strongly and independently associated with both smoking and with increased CHD risk (Albus *et al.*, 2004). Depression is associated with increased levels of coagulation factors VII and X (Doulalas *et al.*, 2005) and also with smoking (e.g., Kavanaugh *et al.*, 2005 for mothers in the United States). Low socioeconomic status (SES) indicators are strongly associated with increased risk of CHD, as well as with smoking; however, smoking does not appear to explain away the causal relation between low SES over time and increased CHD risks (Lawlor *et al.*, 2005a). Possible biological mechanisms suggested for the low SES-increased CHD risk relations include increased levels of inflammatory cytokines and plasma fibrinogen levels among lower-SES subjects (Steptoe *et al.*, 2002, 2003). However, adjusting for life course socioeconomic position attenuates the association between both fibrinogen and C-reactive protein and CHD risk in British women, but not the relation between smoking and CHD risk, suggesting that the latter but not the former may be causes of increased CHD risk (Lawlor *et al.*, 2005b)

Among CHD patients, low SES is associated with both increased levels of high-sensitivity C-reactive protein (a marker of systemic inflammation) and poorer health outcomes (Lubbock *et al.*, 2005).

Many commentators (e.g., Law *et al.*, 1997, Law and Wald, 2003; Pechacek and Babb, 2004) conclude that confounding can account for at most only a small portion of the association between smoking and CHD risk. Others have suggested that the contribution may be much larger and that, in conjunction with modeling biases, incompletely controlled confounding may explain away most or all of the commonly reported associations between ETS exposures and CHD risk (e.g., Gori, 1995; Nilsson, 2001). As in other areas of epidemiology (Greenland, 2005), different treatments of confounding and modeling biases may help to explain significant differences in findings.

Measurement Errors and Misclassification Biases

Survey subjects often round their responses to convenient numbers, such as multiples of 5 or 10, and even these rounded numbers sometimes reflect wishful thinking. The gap between reported and true exposure numbers (or between reported and true exposure categories, in categorical data analyses) may induce large, systematic biases in the results of statistical analyses that ignore such measurement errors and misclassification errors (Hofler, 2005; Fox *et al.*, 2005).

Which is more likely: that a smoker of 8 cigarettes per day will round down to 5 or up to 10 in answering survey questions? Is misreporting 10 cigarettes per day as 5, or misreporting 5 cigarettes per day as 10, the more common error? The answers to such questions may affect the shapes of dose-response relations estimated from self-reported exposure data. If there is a systematic bias toward under-estimating reported cigarette consumption rates, then effects from higher smoking rates will tend to be attributed to lower levels of smoking.

In much past literature on smoking exposure and CHD risks, parametric or semi-parametric statistical models (such as logistic regression or proportional hazards models, respectively) have been used to estimate dose-response relations and associations *without* explicitly modeling effects of exposure estimation and reporting errors. This can introduce large biases (in either direction) into estimated associations and effects (Luan *et al.*, 2005; Hu *et al.*, 1998). It tends to increase the rate of false-positive findings for associations (e.g., between low levels of smoking and CHD risks) by artificially narrowing confidence intervals. Both individual studies and meta-analyses based on them are subject to biases and inflation of false positives when exposure estimation errors are ignored unless appropriate sensitivity analyses and/or corrections are performed (Fox *et al.*, 2005).

CONCLUSIONS

This paper has examined data on the relation between low levels of active smoking and risk of coronary heart disease, with attention to whether existing data conclusively refute the general hypothesis of hormesis for smoking and CHD risk. A review of previous epidemiological literature on active smoking and CHD risks indicates that smokers of relatively few (e.g., less than 10) cigarettes per day do not appear to suffer significantly increased CHD risks, and in some cases may even have smaller CHD risks than non-smokers. Interpretation of the data is complicated by the finding that smoking is significantly associated with other behaviors (e.g., alcohol and coffee consumption) that have been reported as having J-shaped or U-shaped relations with CHD risks (Kleemola *et al.*, 2000; Murray *et al.*, 2002.)

The papers on active and passive smoking that we have reviewed, including meta-analyses and review papers as well as individual studies (see e.g., Table 1), generally have not corrected systematically and thoroughly for potential biases due to model form selection, variable selection, variable coding, confounding, and errors in exposure estimates, although appropriate statistical methods for doing so have been developed (Ricci and Cox, 2002; Greenland, 2005). Such uncorrected biases may help to explain differences in conclusions and interpretations of data across studies.

To avoid potential biases due to model form selection, we examined the shape of the dose-response relation for active smoking and CHD risks in the NHANES data set, without using parametric models to smooth (and perhaps distort) response data at the lowest exposure levels using data collected at higher exposure levels. The major finding, shown in Figures 1 and 2, is that the empirical relations we found in this data set do *not* appear to refute the possibility of hormesis. Nor do they support the reverse hypothesis, that low levels of exposure are disproportionately hazardous compared to higher levels of exposure (super-linear dose-response). This is partly consistent with a recent large multinational study of acute myocardial infarction risk (Teo *et al.*, 2006, Table 4) that found odds ratios increasing approximately linearly with cigarettes per day (with statistically significant increases noted at reported smoking levels as low as 3-4 cigarettes per day). However, for the dependent variable examined in this paper (answering "Yes" to the question "Has a doctor or other health professional ever told you [or subject] that {you/s/he} had coronary heart disease?"), no significant increase in risk is apparent at the lowest levels of smoking (Fig. 1).

Finally, we reviewed relevant biological literature and noted evidence of U-shaped and J-shaped relations between several important smoking-related risk factors (and/or biomarkers) and CHD risks (e.g., Bonora *et al.*, 1998; Curb *et al.*, 2004; Dyer *et al.*, 1980; El-Khairy *et al.*, 2001; Flack *et*

al., 1995, Gagnon *et al.*, 1994; Iribarren *et al.*, 1996). Such studies suggest, but do not prove, that the appearance of reduced risks at low levels of smoking exposure may be caused in whole or in part by confounding due to relatively low levels of other risk factors at these low levels of smoking.

ACKNOWLEDGEMENT

The work reported was supported in part by Philip Morris International (PMI). The opinions, data examined, and research questions addressed (especially, whether the hypothesis of hormesis is inconsistent with NHANES data) are solely the author's, who also retained the right to publish all findings without review by PMI. I am grateful to Dr. Edward Sanders and Professor Paolo Ricci for stimulating, substantive discussions on modeling challenges in CHD epidemiology.

REFERENCES

Albus C, Jordan J, Herrmann-Lingen C. 2004. Screening for psychosocial risk factors in patients with coronary heart disease-recommendations for clinical practice. Eur J Cardiovasc Prev Rehabil. Feb;11(1):75-9.

Austin PC, Tu JV. 2004. Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality. J Clin Epidemiol. Nov;57(11):1138-46.

Barnoya J, Glantz SA. 2005. Cardiovascular effects of secondhand smoke: nearly as large as smoking. Circulation. May 24;111(20):2684-98.

Bonora E, Willeit J, Kiechl S, Oberhollenzer F, Egger G, Bonadonna R, Muggeo M. 1998. U-shaped and J-shaped relationships between serum insulin and coronary heart disease in the general population. The Bruneck Study. Diabetes Care. Feb;21(2):221-30.

Brookes ST, Whitley E, Peters TJ, Mulheran PA, Egger M, Davey Smith G. 2001. Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives. Health Technol Assess.5(33):1-56.

Bush TL, Comstock GW. Smoking and cardiovascular mortality in women. 1983. Am J Epidemiol. Oct;118(4):480-8.

Calabrese EJ, Baldwin LA. 2001. The frequency of U-shaped dose responses in the toxicological literature. Toxicol Sci. Aug;62(2):330-8.

Chen R, Tavendale R, Tunstall-Pedoe H. 2004. Environmental tobacco smoke and prevalent coronary heart disease among never smokers in the Scottish MONICA surveys. Occup Environ Med. Sep;61(9):790-2.

Chien KL, Sung FC, Hsu HC, Su TC, Chang WD, Lee YT. 2005. Relative importance of atherosclerotic risk factors for coronary heart disease in Taiwan. Eur J Cardiovasc Prev Rehabil. Apr;12(2):95-101.

Curb JD, Abbott RD, Rodriguez BL, Masaki K, Popper J, Chen R, Petrovitch H, Blanchette P, Schatz I, Yano K. 2004. Prospective association between low and high total and low-density lipoprotein cholesterol and coronary heart disease in elderly men. J Am Geriatr Soc. Dec;52(12):1975-80.

Doulalas AD, Rallidis LS, Gialernios T, Moschonas DN, Kougioulis MN, Rizos I, Tselegaridis TS, Kremastinos DT. 2005. Association of depressive symptoms with coagulation factors in young healthy individuals. Atherosclerosis. May;186(1):121-5

Dreiseitl S, Ohno-Machado L, Vinterbo S. 1999. Evaluating variable selection methods for diagnosis of myocardial infarction. Proc AMIA Symp.:246-50.

Dyer AR, Persky V, Stamler J, Paul O, Shekelle RB, Berkson DM, Lepper M, Schoenberger JA, Lindberg HA. 1980. Heart rate as a prognostic factor for coronary heart disease and mortality: findings in three Chicago epidemiologic studies. Am J Epidemiol. Dec;112(6):736-49.

Ebrahim S, Montaner D, Lawlor DA. 2004. Clustering of risk factors and social class in childhood and adulthood in British women's heart and health study: cross sectional analysis. BMJ. 2004 March;328(7444):861:1-5. doi:10.1136/bmj.38034.702836.55

El-Khairy L, Ueland PM, Refsum H, Graham IM, Vollset SE; European Concerted Action Project. 2001. Plasma total cysteine as a risk factor for vascular disease: The European Concerted Action Project. Circulation. May 29;103(21):2544-9.

Enstrom JE, Kabat GC. 2003. Environmental tobacco smoke and tobacco related mortality in a prospective study of Californians, 1960-98. BMJ. May 17;326(7398):1057:1-10.

Flack JM, Neaton J, Grimm R Jr, Shih J, Cutler J, Ensrud K, MacMahon S. 1995. Blood pressure and mortality among men with prior myocardial infarction. Multiple Risk Factor Intervention Trial Research Group. Circulation. Nov 1;92(9):2437-45.

Fox MP, Lash TL, Greenland S. 2005. A method to automate probabilistic sensitivity analyses of misclassified binary variables. Int J Epidemiol. Dec;34(6):1370-6.

Freund KM, Belanger AJ, D'Agostino RB, Kannel WB. 1993. The health risks of smoking. The Framingham Study: 34 years of follow-up. Ann Epidemiol. Jul;3(4):417-24.

Gagnon DR, Zhang TJ, Brand FN, Kannel WB. 1994. Hematocrit and the risk of cardiovascular disease — the Framingham study: a 34-year follow-up. Am Heart J. Mar;127(3):674-82.

Glantz SA, Parmley WW. 1991. Passive smoking and heart disease. Epidemiology, physiology, and biochemistry. Circulation. Jan;83(1):1-12.

Gori GB. 1995. Environmental tobacco smoke and coronary heart syndromes: Absence of an association. Regul Toxicol Pharmacol. Apr;21(2):281-95.

Greenland S. 2005. Multiple-bias modelling for analysis of observational data. J Roy Stat Soc. 168:267-291.

Hatsukami DK, Kotlyar M, Allen S, Jensen J, Li S, Le C, Murphy S. 2005. Effects of cigarette reduction on cardiovascular risk factors and subjective measures. Chest. Oct;128(4):2528-37.

Hofler M. The effect of misclassification on the estimation of association: a review. 2005. Int J Methods Psychiatr Res. 4(2):92-101.

Iribarren C, Sharp D, Burchfiel CM, Sun P, Dwyer JH. 1996. Association of serum total cholesterol with coronary disease and all-cause mortality: multivariate correction for bias due to measurement error. Am J Epidemiol. Mar 1;143(5):463-71.

Hu P, Tsiatis AA, Davidian M. 1998. Estimating the parameters in the Cox model when covariate variables are measured with error. Biometrics. Dec;54(4):1407-19.

Jenkins CD, Rosenman RH, Zyzanski SJ. 1968. Cigarette smoking. Its relationship to coronary heart disease and related risk factors in the Western Collaborative Group Study. Circulation. Dec;38(6):1140-55.

Kavanaugh M, McMillen RC, Pascoe JM, Hill Southward L, Winickoff JP, Weitzman M. 2005. The cooccurrence of maternal depressive symptoms and smoking in a national survey of mothers. Ambul Pediatr. 2Nov-Dec;5(6):341-8.

Kleemola P, Jousilahti P, Pietinen P, Vartiainen E, Tuomilehto J. 2000. Coffee consumption and the risk of coronary heart disease and death. Arch Intern Med. Dec 11-25;160(22):3393-400.

Law MR, Morris JK, Wald NJ. 1997. Environmental tobacco smoke exposure and ischaemic heart disease: an evaluation of the evidence. BMJ. Oct 18;315(7114):973-80.

Law MR, Wald NJ. 2003. Environmental tobacco smoke and ischemic heart disease. Prog Cardiovasc Dis 46:31-8.)

Lawlor DA, Ebrahim S, Davey Smith G. 2005a. Adverse socioeconomic position across the lifecourse increases coronary heart disease risk cumulatively: findings from the British women's heart and health study. J Epidemiol Community Health. Sep;59(9):785-93.

Lawlor DA, Smith GD, Rumley A, Lowe GD, Ebrahim S. 2005b. Associations of fibrinogen and C-reactive protein with prevalent and incident coronary heart disease are attenuated by adjustment for confounding factors. British Women's Heart and Health Study. Thromb Haemost. May;93(5):955-63.

Lee J, Heng D, Chia KS, Chew SK, Tan BY, Hughes K. 2001. Risk factors and incident coronary heart disease in Chinese, Malay and Asian Indian males: the Singapore Cardiovascular Cohort Study. Int J Epidemiol. Oct;30(5):983-8.

Luan X, Pan W, Gerberich SG, Carlin BP. 2005. Does it always help to adjust for misclassification of a binary outcome in logistic regression? Stat Med. Jul 30;24(14):2221-34.

Lubbock LA, Goh A, Ali S, Ritchie J, Whooley MA. 2005. Relation of Low Socioeconomic Status to C-Reactive Protein in Patients With Coronary Heart Disease (from the Heart and Soul Study). Am J Cardiol. Dec 1;96(11):1506-11.

MacCallum PK. 2005. Markers of hemostasis and systemic inflammation in heart disease and atherosclerosis in smokers. Proc Am Thorac Soc. 2(1):34-43.

Miller M, Zhan M, Havas S. 2005. High attributable risk of elevated C-reactive protein level to conventional coronary heart disease risk factors: the Third National Health and Nutrition Examination Survey. Arch Intern Med. Oct 10;165(18):2063-8.

Murray RP, Connett JE, Tyas SL, Bond R, Ekuma O, Silversides CK, Barnes GE. 2002. Alcohol volume, drinking pattern, and cardiovascular disease morbidity and mortality: is there a U-shaped function? Am J Epidemiol. Feb 1;155(3):242-8.

Nilsson R. 2001. Environmental tobacco smoke revisited: the reliability of the data used for risk assessment. Risk Anal. Aug;21(4):737-60.

Njolstad I, Arnesen E, Lund-Larsen PG. 1996. Smoking, serum lipids, blood pressure, and sex differences in myocardial infarction. A 12-year follow-up of the Finnmark Study. Circulation. Feb 1;93(3):450-6.

O'callaghan PA, Fitzgerald A, Fogarty J, Gaffney P, Hanbidge M, Boran G, Enright H, Murphy J, McCarthy B, Graham IM. 2005. New and old cardiovascular risk factors: C-reactive protein, homocysteine, cysteine and von Willebrand factor increase risk, especially in smokers. Eur J Cardiovasc Prev Rehabil. Dec;12(6):542-547.

Pechacek TF, Babb S. 2004. How acute and reversible are the cardiovascular risks of secondhand smoke? BMJ. Apr 24;328(7446):980-3.

Ricci, PF and Cox, LA 2002. Empirical causation and biases in epidemiology: Issues and solutions. Technology, 9:23-53. http://www.cognizantcommunication.com/filecabinet/Technology/tech91abs.html#tech91abs3

Royston P, Altman DG, Sauerbrei W. 2006. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med. Jan 15;25(1):127-41.

Steenland K, Thun M, Lally C, Heath C Jr. 1996. Environmental tobacco smoke and coronary heart disease in the American Cancer Society CPS-II cohort. Circulation. Aug 15;94(4):622-8.

Steptoe A, Owen N, Kunz-Ebrecht S, Mohamed-Ali V. 2002. Inflammatory cytokines, socioeconomic status, and acute stress responsivity. Brain Behav Immun. Dec;16(6):774-84.

Steptoe A, Kunz-Ebrecht S, Owen N, Feldman PJ, Rumley A, Lowe GD, Marmot M. 2003. Influence of socioeconomic status and job control on plasma fibrinogen responses to acute mental stress. Psychosom Med. Jan-Feb;65(1):137-44.

Steyerberg EW, Eijkemans MJ, Boersma E, Habbema JD. 2004. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. Stat Med. Aug 30;23(16):2567-86.

Stromberg U. 1996. Collapsing ordered outcome categories: a note of concern. Am J Epidemiol. Aug 15;144(4):421-4.

Teo KK, Ounpuu S, Hawken S, Pandey MR, Valentin V, Hunt D, Diaz R, Rashed W, Freeman R, Jiang L, Zhang X, Yusuf S; INTERHEART Study Investigators. 2006.Tobacco use and risk of myocardial infarction in 52 countries in the INTERHEART study: a case-control study. Lancet. Aug 19;368(9536):647-58.

Wells AJ. 1994. Passive smoking as a cause of heart disease. J Am Coll Cardiol. Aug;24(2):546-54