
Demonstration of Deductive Meta-Analysis: Ethanol Intake and Risk of Myocardial Infarction

Malcolm Maclure

INTRODUCTION

Meta-analyses should be analytic and deductive. In a review of the state of the science of meta-analysis in the previous volume of *Epidemiologic Reviews*, a list of definitions and synonyms of meta-analysis was given: "overview, pooling, data pooling, literature synthesis, data synthesis, quantitative synthesis, and quantitative review" (1, p. 154). Indeed, most meta-analyses are more synthetic than analytic: They produce a summary, such as an aggregate relative risk and 95 percent confidence interval, from a set of individual studies and stop there. Such a "meta-synthesis" has an inductive approach, i.e., generalization from a set of particular observations. By contrast, a deductive approach starts with alternative generalizations (hypotheses) and uses particular observations to discriminate among them. The causal hypothesis of primary interest is considered corroborated if competing hypotheses do not stand up to the evidence (2).

The words "inductive" and "deductive" here have the meanings used in logic (3). Induction is the inference "If true for A, then true for B" when A is a part, sample, or special case of B. When epidemiologists infer

that a relation probably holds in the general population because it holds in a study population, they are using inductive inference. Deduction is the opposite—to infer "If not true for A, then not in general true for B" when A is part of B. When epidemiologists use likelihoods in calculations, i.e., hypothesize that a relation holds in the general population and then ask how probable the study data are given this hypothesis, they are using deductive inference.

The need for a deductive (refutationist, Popperian) approach to epidemiology has been asserted (2, 4–7) and disputed (7–9) but rarely demonstrated using concrete examples (2). This review provides such a demonstration. By example more than theoretical arguments, it illustrates a refutationist principle: *Like differential diagnosis of ambiguous symptoms, causal inference proceeds by the deductive process of ruling out noncausal alternative explanations.* The example presented is a meta-analysis of epidemiologic evidence for and against the existence of a causal relation between ethanol intake and risk of myocardial infarction in particular and coronary heart disease in general.

Dickersin and Berlin (1) stressed that a meta-analysis should go beyond weighted averaging of several studies' results and include analyses aimed at explaining inconsistencies. I would go further and recommend that such analyses include ancillary data relevant to competing hypotheses and be structured using deductive reasoning, as shown below. In an earlier volume of *Epidemiologic Reviews*, Greenland concluded, "Meta-analytic and narrative (qualitative)

Received for publication July 1, 1992, and in final form July 22, 1993.

Abbreviations: CI, confidence interval; HDL, high density lipoprotein cholesterol; RR, relative risk.

From the Department of Epidemiology, Harvard School of Public Health, Boston, MA, and the Institute for Prevention of Cardiovascular Disease, New England Deaconess Hospital, Boston, MA.

Reprint requests to Dr. Malcolm Maclure at the Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115.

aspects of research review can and should be complementary ... [C]ausal explanation of similarities and differences among study results noted in a meta-analysis is a qualitative aspect of the review, and thus outside the realm of statistical meta-analysis" (10, p. 28). He used the word "outside" heuristically. Here I show how the qualitative and quantitative realms are intricately interwoven.

HYPOTHESES

Meta-analyses, like all research studies, should begin with competing hypotheses. Data do not speak for themselves. Without at least rudimentary hypotheses, observation itself is not possible. Precise prior hypotheses improve an investigator's powers of observation.

This review begins by articulating alternative hypotheses for how and why ethanol intake may be related to incidence of myocardial infarction. Fatal coronary disease and mixed coronary outcomes are considered as proxies (with lower specificity) for the purer outcome, infarction. We are not concerned here with total coronary disease or total mortality, for which a whole series of additional hypotheses would need to be considered.

Hypotheses concerning the relation of ethanol intake to risk of infarction include: 1) the null hypothesis, that the dose-response relation is flat; 2) the causative hypothesis, that ethanol is a risk factor; 3) the preventive hypothesis, that ethanol intake reduces risk; and 4) a large set of "distortion hypotheses": publication bias, selection bias, outcome misclassification bias, exposure misclassification bias, confounding, or reverse causation (the outcome influences the probability of exposure).

Many combinations of these hypotheses are possible, as figure 1 illustrates. For example, in the third row and second column of figure 1, the "U-hypothesis" (11) is illustrated, i.e., the hypothesis that ethanol is preventive at low doses and causative at high doses.

The most popular distortion hypotheses are that the higher incidence of infarction among teetotalers may be due to contamination by: 1) former moderate drinkers who decided to quit when they became sick, because they believed alcohol to be a risk factor (reverse causation); 2) current heavy drinkers who lie about their intake (pure exposure misclassification); 3) former heavy drinkers, many of whom are former or current heavy smokers (exposure misclassification and confounding); and 4) nondrinkers

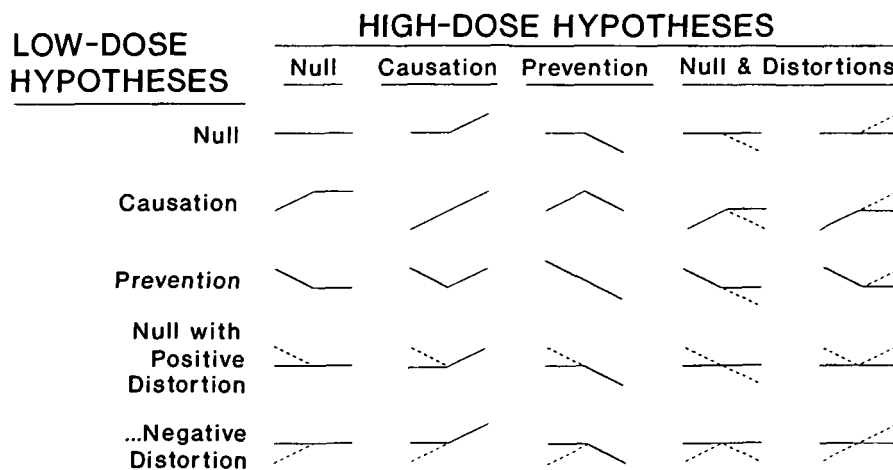


FIGURE 1. Array of competing hypotheses about the shape of the dose-response relation between ethanol intake and risk of coronary heart disease. Causal relations are shown by solid lines, and distortions (deviations of the empirical association from the underlying causal relation) are shown as broken lines.

who are more sedentary and obese than the average drinker, or who get more of their dietary energy from fats than the average drinker (pure confounding).

Among the consequences of beginning a meta-analysis with a set of hypotheses are: 1) a broader search of the literature for evidence, preferably meta-analyses, concerning competing hypotheses (e.g., the relations of ethanol intake to Type A personality and of Type A personality to heart disease); 2) greater alertness to such evidence in reports from studies of the causal relation of interest (e.g., calculation of an estimate of selection bias by comparing components of a hospital control group); and 3) categorization of studies by more study characteristics pertinent to tests of competing hypotheses in meta-regression (e.g., whether ex-drinkers were excluded from the nondrinker group).

METHODS

Like a deductive approach to multivariate analysis (2), a deductive meta-analysis presents not a single conclusion but a description of the survival or refutation of hypotheses over a course of tests against data and against each other. From the first stage of selecting evidence to the final sensitivity analyses and writing of the report, the goal is to discriminate among competing hypotheses.

Selection of studies

An important competing hypothesis is that publication bias has distorted the overall relation between ethanol and infarction in the total body of published data. Tracking down all unpublished data may be the best way of avoiding publication bias when meta-analyzing randomized controlled trials, but with nonexperimental studies, there is no clear criterion analogous to "all patients randomized." "All patients studied" would include any cohort, case-control, cross-sectional, or even ecologic study that collected data on alcoholic drinks and heart disease. Thus, including all published and

retrievable unpublished reports may make a meta-analysis of nonexperimental data more, rather than less, susceptible to publication bias. This is a special case of the well known trade-off between precision and bias, i.e., between quantity and quality of data.

To refute publication bias, sometimes it may be better to restrict the meta-analysis to study populations that are relatively immune to the bias. In the case of ethanol and infarction, recent reports giving data on multiple covariates from large, well known cohorts with ongoing funding are probably relatively immune compared with the average case-control study. However, even investigators of well known cohorts are apt to delay publication until an association "achieves statistical significance."

The present meta-analysis aimed for an initially inclusive selection of studies, but in testing of competing hypotheses, the selection was inevitably whittled down to those providing the best data, i.e., those most resistant to uncertainty concerning competing hypotheses. Studies published between 1968 and early 1993 were identified by a MEDLINE® search (National Library of Medicine) and were supplemented with additional references cited in previous reviews (11, 12). This review focused on data from 42 reports (13–54). Twenty-seven reports (55–81) were immediately excluded, mainly because of overlapping data, but also because of weaknesses: use of prevalent cases (72, 78), crude exposure groupings (58, 63, 64, 76, 81), a combination of mixed outcomes and small numbers (66, 69, 75), or limited analyses (55, 60, 61). Criteria for choosing which reports to include when there were multiple reports from the same cohort were (in order of priority):

1. Nonfatal myocardial infarction was the outcome, because fatal and mixed outcomes have lower specificity.
2. Ethanol intake was categorized in more detail, because this enabled better assessment of the shape of the dose-response curve.
3. Cases were more numerous.

Data from some of the omitted reports were used in subsidiary analyses. Data related to competing hypotheses were sought from sources other than studies of ethanol and heart disease but less systematically, a weakness discussed below.

Data extraction and unification

For testing of exposure misclassification bias, some studies (58, 63, 64, 76, 81) were excluded because of inadequate exposure assessment. For example, the Walnut Creek Contraceptive Drug Study (58) classified subjects only as drinkers or nondrinkers, and the study of Italian rural men (81) had, as the lowest category of intake, 40 g/day or less of ethanol.

The quality of exposure information in the included studies varied greatly. All studies used self-reported beverage consumption frequencies, but some (16, 19, 20, 22, 38, 53) asked only one question about alcoholic beverages, rather than separate questions for beer, wine, and liquor. There were large differences in time periods to which the questions referred: the 24-hour period before the interview (26), the past 2–3 days (45), the past week (39, 49), the past month (30, 36, 52), the past year (22, 27–29, 51), and “usual” frequency (table 1). The number of levels of intake into which subjects were classified ranged from two to more than 10.

A level of ethanol intake was assigned to each relative risk using the following algorithm:

1. If the mean intake of alcoholic drinks within a subgroup was provided or could be estimated from a table or graph of the distribution, the only conversion made was to express intake in grams of ethanol per day, assuming that one drink contains 13 g of ethanol.
2. If the paper defined subgroups using contiguous intervals of ethanol levels (e.g., 0.1–1.0, 1.1–29, and ≥ 30 ounces/month), the cutpoints were converted to grams of ethanol per day and the mean intake in each interval was estimated, assuming that the study had the same distribution of alcohol intake as the National Health Interview Survey (82).
3. If subgroups were defined by the most common values (e.g., <1 , 1–2, or ≥ 3 drinks/day), then the scale was converted to a set of contiguous intervals, picking logical cutpoints between the common values (e.g., 0.75 drinks/day as the cutpoint between <1 and 1–2 drinks/day); the cutpoints were converted to grams of ethanol per day, and the mean for each interval was estimated as indicated above.
4. When subgroups were defined in more unusual ways, these were translated into contiguous intervals as well as possible, and the above steps were implemented.

For testing the hypothesis that the dose-response curve was distorted by error in this algorithm, sensitivity of the results to exposure classification was assessed by changing assumptions and repeating the regressions.

Quadratic meta-regression

To test competing hypotheses fairly, we need a method of meta-regression that does not force one shape on the data. In the Albany Study (47), for example, the regression forced a straight line and obscured a U-shape apparent in the crude data. Quadratic regression accommodates many more shapes than does linear regression. Recently a method (83) was described for quadratic meta-regression of dose-response data, which takes account of the fact that dose-specific relative risks are never independent. (Their interdependence arises from sharing the same reference group, and results in double—or multiple—counting of the uncertainty of the denominator of each ratio.) The one-step “pool-first” method was used: The dose-specific confounder-adjusted logarithms of the relative risks from all studies were pooled, and a curve was fitted by weighted quadratic regression. The weights were the inverses of the covariance-adjusted

TABLE 1. Characteristics of cohort studies (1968–1993) of ethanol intake and fatal or nonfatal coronary heart disease that were included in a meta-analysis that suggested an L-shaped dose-response relation

Cohort study (in chronologic order)	Exposure assessment			Outcomes			
	Instrument*	Time†	No. of levels‡	Years of follow-up	Types§	No. of cases	RR
Honolulu Heart Study (23, 24)	I	Ever	4	6	NF	278	0.61
Chicago Western Electric Company Study (33)	E	Current	5	17	F	149	1.8
Yugoslavia Cardiovascular Disease Study (25, 34)	I	<40 g/month	4	7	NF	166	0.40
Kaiser Permanente matched cohorts study (35)	E	1 year	4	10	F	8,060	0.70
Whitehall Study (45)	I	3 days	3	10	M	63	0.59
Puerto Rico Heart Health Program (26)	I	24 hours	4	8	NM	170	0.58
North Karelia Project (eastern Finland) (36)	E	1 year	3	7	F	95	1.16
Framingham Study (46)	E	Current	8	22	M	532	0.72
Study of Massachusetts elderly (37)	I	Current	3	4.8	F	42	0.58
Study of Japanese physicians (38)	Q	20 years	4	19	F	123	0.64
Lipid Research Clinics Follow-up Study (39)	E	1 week	4	8.5	F	89	0.82
British Regional Heart Study (48)	I	Ever	7	7.5	M	265	1.02
Albany Study (47)	E	Current	6	18	M	348	0.84
Albany Study (47)	E	Current	6	10	M	159	0.77
Finnish Mobile Clinic Health Survey (40)	I	Current	2	5	F	140	1.5
Nurses' Health Study (28)	Q	1 year	5	4	NF	200	0.55
American Cancer Society Prospective Study (women) (32)	Q	Current	8	12	F	19,661	0.94
Kaiser Permanente cohort study (27)	E	Ever	7	2.5	N	355	0.66
Nutrition Canada Survey Cohort Study (44)	I	Current	2	10	M	155	0.90
Finnish rural cohorts of the Seven Countries Study (30)	E	30 days	2	10	F	119	1.1
St. James Survey (Trinidad) follow-up (49)	I	1 week	4	10	M	49	0.46
NHANES I¶ Epidemiologic Follow-up Study (43)	E	Current	>9	10	F	287	1.0
American Cancer Society Prospective Study (men) (31)	Q	Current	8	12	F	18,771	0.83
Health Professionals Follow-up Study (29)	Q	1 year	6	1.5	NFM	350	0.64
MRFIT¶¶ (50)	I	Current	4	10	F	427	0.88
Alameda County Study (42)	Q	Current	2	10	F	135	0.65
Normative Aging Study (51)	Q	1 year	3	12	F	74	0.63
Established Populations for Epidemiologic Study of the Elderly (52)	I	1 year	3	5	M	677	0.73
Busselton Population Study (53)	Q	Current	3	13	F	325	0.66
Copenhagen Male Study (54)	Q	Current	5	4	M	100	0.73

* Instrument used to measure ethanol intake. Q, questionnaire; I, interview; E, either Q or I in connection with a physical examination.

† Time period defining "nondrinker"; i.e., the period to which questions on alcohol intake referred.

‡ No. of levels into which alcohol drinkers were categorized.

§ Types of outcome: N, nonfatal myocardial infarction; F, fatal coronary heart disease; M, a mixture of nonfatal and fatal coronary outcomes.

|| RR, relative risk of coronary disease, a weighted average of all relative risks for ethanol doses of ≥ 10 g/day.

¶ NHANES I, First National Health and Nutrition Examination Survey; MRFIT, Multiple Risk Factor Intervention Trial.

variances of the logarithms of the relative risks. Quadratic regression analyses of the square root of ethanol intake were also conducted.

For testing the hypothesis that contamination of the nondrinking reference group was distorting the shape of the curve, a similar regression was carried out with an indi-

cator term for studies that did not separate ex-drinkers from never drinkers.

For testing the preliminary global hypothesis that one or more biases or effect modification, rather than random sampling variation, accounted for differences among studies, heterogeneity was assessed by computing a goodness-of-fit statistic, Δ , the root

weighted-mean square of the residual deviation (on the logarithmic scale) of each observed relative risk (RR_i) from the corresponding relative risk predicted by the regression (RR_{ip}):

$$\Delta = \left\{ \left[\sum_{i=1}^n w_i (\ln RR_i - \ln RR_{ip})^2 \right] / (n - t) \right\}^{0.5},$$

where n is the number of relative risks, w_i is the weight given to each RR (the covariance-adjusted inverse variance of the logarithm of RR_i), and t is the number of terms in the regression model (84, 85). The summation in the numerator of Δ has a chi-squared distribution with $n - t$ degrees of freedom. If each study population were a perfect random sample from a single common source population, the expected value for Δ would be about 1.0. If the observed Δ is significantly greater than 1.0, it represents "beyond-random" variation (i.e., heterogeneity) among relative risks, and serves as a crude estimate of the net systematic variation caused by bias or effect modification. Possible causes of heterogeneity were explored by deleting individual studies that appeared to be outliers, by deleting certain types of study, or by repeating the regression with indicator terms for study flaws and characteristics.

For testing the hypotheses that weaker study designs and nonspecific outcome measures were distorting results, separate quadratic meta-regressions were carried out for: 1) case-control studies of nonfatal infarction (13–18); 2) case-control studies of coronary deaths (19–22); 3) case-control studies with population controls, combining both nonfatal and fatal cases (18–22); 4) cohort studies with nonfatal infarction as the outcome (23–29); 5) cohort studies with coronary death as the outcome (28–30, 33–41, 43, 50, 51, 53, 54), excluding the huge American Cancer Society studies (31, 32), which had more cases than all other cohorts combined; 6) cohort studies with other or mixed outcomes: coronary artery bypass grafts or percutaneous transluminal coronary angioplasty (29); a combination of nonfatal and fatal heart disease, including an-

gina (26, 46, 47, 49); coronary insufficiency (26); or cardiovascular death including stroke (44, 45, 48, 52); and 7) cohort studies of women with any cardiovascular outcome (27, 28, 32, 35, 37, 39, 41–44, 46, 53, 58).

Incremental relative risks and regression

Even a quadratic model may distort a dose-response relation: A kink in the true curve may be smoothed over. Given enough data, therefore, shape is better assessed using nonparametric methods; for example, incremental relative risks (86) or moving line regression (87).

Dose-specific relative risks from all cohort studies were pooled for an incremental regression, a variation on moving line regression, as follows. The first step was a regular weighted linear regression between 0 and 2.0 g/day of ethanol. No intercept term was used, thus forcing the line through the origin. A meaningful central point on this first regression line—0.8 g/day (about two standard drinks per month), at which the relative risk was 0.95—was taken as the origin for the next weighted regression, in the interval 0.8–5.0 g/day. At a meaningful central point in this next interval—3.8 g/day (about two drinks per week), the coefficient represents the logarithm of the incremental relative risk for an increase from two drinks per month to two drinks per week. By repeating this process, the dose-response relation was described for dose increments from two drinks per week to four drinks per week to one, two, five, and about 10 drinks per day. Similar incremental regressions were carried out with varying bandwidths. After an L-shape was found (i.e., a drop in risk at low doses, followed by a plateau in risk above one drink per day), overall non-incremental regressions were carried out assuming an L-shape.

For retesting the hypothesis that higher incidence in nondrinkers was due to contamination of the nondrinkers with ex-drinkers, incremental analyses were repeated in the low-dose range, limited to data from studies that separated ex-drinkers from

nondrinkers and had a category of occasional (less than daily) drinkers (23, 27, 35, 38, 42), and by including in the meta-regression an indicator term for studies that did not separate ex-drinkers from nondrinkers.

Quantifying bias

Rather than addressing the possibility of confounding and misclassification bias simply qualitatively, the magnitude of confounding was estimated by computing a relative risk ratio where possible. For example, the crude relative risk before age adjustment was divided by the age-adjusted relative risk to obtain the relative risk ratio for age. This ratio represents the spurious relative risk that would have been observed if age were the only confounder and ethanol had no causative or preventive effect. It is an estimate of the parameter U , which quantifies confounding (10). Greenland advocates adjusting relative risks using these ratios (10, 88). Such adjustments should be treated as sensitivity analyses, and ideally should incorporate information on uncertainty of the ratio estimates. In this meta-analysis, such adjustments were not made, because with so many good large cohort studies, it was possible to analyze separately in the regression those studies with adequate internal adjustments. The relative risk ratio was also used to quantify possible bias due to non-differential misclassification.

Inclusion of study characteristics in regressions

For testing the hypothesis that a particular combination of study flaws might account for the association, multivariate meta-regressions were performed. Indicator terms were assigned to individual relative risks or studies, indicating the presence or absence of a study flaw (e.g., contamination of nondrinkers with ex-drinkers, a nonspecific outcome measure, no control of confounding by age). When the study characteristic could be quantified, ordinal or continuous variables were assigned (e.g., number of exposure categories, number of years of follow-up). The model was fitted by "forward

elimination" (2), i.e., starting with the crude association of interest and adding terms for potential biases to see which combinations of study flaws and characteristics might explain the association, based on the change-in-estimate criterion (89).

Some authors use additional quality weighting according to a systematic assessment of study quality by a panel of reviewers (90). This can be viewed as analogous to expanding the confidence interval of each relative risk to reflect the panel's judgment of the amount of additional uncertainty due to departures from perfect study quality. The weaknesses of this approach are that it attributes no directionality to the uncertainty (the confidence interval is widened equally in both directions) and it assigns a magnitude of uncertainty based more on personal judgment than on the data at hand. By contrast, multivariate meta-regression of relative risks on study characteristics allows estimation of the directionality and magnitude of bias (i.e., the relative risk ratio) attributable to each hypothesized flaw, conditional on other hypothesized flaws. Thus, multiple hypotheses about distortions can be tested.

RESULTS

Refutation of publication bias

In his meta-analysis of coffee intake and coronary risk (89), Greenland cited 14 cohort studies, seven (91–97) of which did not contribute to the present meta-analysis because I found no published reports from them concerning ethanol intake. The seven cohorts yielded a total of 2,378 coronary outcomes, which compares with 38,432 cases from the American Cancer Society studies and 19,413 from the other cohorts included in this review. Let us assume that all of these cohort studies had data on ethanol intake and, to be conservative, that the association in these studies between ethanol and coronary disease was positive (i.e., in the opposite direction from what is reported here). Suppose the average relative risk in these other cohorts were 1.2, because if it had been greater we would expect to have seen reports of significant positive associa-

tions. This would amount to a hypothetical relative risk ratio of 1.02 for publication bias due to exclusion of these cohorts. Such a weak bias would have negligible influence on our conclusions.

The relation of ethanol intake to risk of heart disease has been a prominent controversial topic for long enough that null findings should now be as interesting and publishable as non-null results. In the present review, therefore, publication bias was considered reduced in analyses that effectively were restricted to recently reported cohort studies. It turned out that controlling for study quality in the meta-regressions effectively controlled for recency of publication, because the quality of design and analysis of epidemiologic studies has improved over the decades. The best studies are the recently published large cohort studies that use multivariate methods.

On the basis of this evidence and the quality of excluded reports, I hypothesize that exclusion of unpublished and inadequately analyzed data from the meta-regression probably did not produce much bias. Unfortunately, the alternative explanation, that publication bias is a major cause of the observed association, will remain unrefuted until further empirical studies of publication bias are carried out (98, 99).

Publication bias also operates against evidence concerning competing hypotheses. Analyses of selection bias and misclassification bias at low doses of alcohol consumption are hard to find (although there is a great deal of literature concerning heavy drinkers). Although I found some meta-analyses concerning potential confounders (100–103), a more thorough search of the literature for ancillary data concerning hypothesized confounders was not within the resources available. Even within the studies of ethanol and infarction themselves, there was plenty of “semi-publication” bias; most reports gave only glimpses of the impact of adjusting for covariates.

Refutation of selection bias

It was not possible to refute selection bias or recall bias as an explanation for the as-

sociations seen in case-control studies. The heterogeneity among relative risks from case-control studies was much greater than would be expected from random sampling variation ($\Delta = 2.4$; $p < 0.001$). Such heterogeneity could be due to bias or effect modification. Repetition of regressions after deletion of outliers indicated that the heterogeneity was not due to a few deviant points or a single deviant study.

Selection bias probably occurred in studies with hospital controls (13–16). In two (13, 16) of those studies, patients with cholecystectomies or trauma were not excluded from the control groups; gallstones and accidents are now known to be related to ethanol intake (104, 105). In two others (14, 15), 40 percent of the controls had been admitted for disc disorders. Both studies found the same elevated ethanol intake among patients with disc disorders as compared with the other hospital controls, equivalent to a relative risk of 1.3. If this finding is real, it would have exaggerated the ethanol-infarction association, corresponding to a relative risk ratio of 0.88 for occasional drinking (table 2). Excluding studies with hospital controls (13–17) only slightly reduced the heterogeneity ($\Delta = 2.2$). Inclusion of indicator terms for women and fatal outcomes in the meta-regression (table 2) further reduced the heterogeneity ($\Delta = 2.0$; $p < 0.001$), but it was still too large.

Although effect modification is a possible explanation for the large remaining heterogeneity, we cannot rule out selection bias, especially because the exposure of interest—alcohol drinking—is believed to be related to sociability and antisocial behavior in complex ways. Therefore, to guard against selection bias, I excluded case-control studies from further meta-analyses. Data from population-based case-control studies are shown in figure 2.

Refutation of chance

Each of the three quadratic meta-regressions of cohort data yielded similar results: a significant inverse association with significant positive curvature (table 3). Quadratic models fitted well to relative risks for

TABLE 2. Estimated relative risk ratios for hypothesized sources of bias and effect modification in case-control studies of ethanol intake and risk of coronary heart disease

Source	RRR*	95% CI*	Studies	Estimation
Hospital control bias	2.1	0.5–8.5	All case-control studies (13–22)	RR for 10 g/day of ethanol from meta-regression of data from hospital case-control studies divided by the same RR for population control studies.
Hospital control bias	0.88	0.77–1.0	Northeastern US studies (14, 15)	$1 + 0.40 \times (RR - 1)$, where RR is the odds ratio for drinkers vs. nondrinkers, comparing controls with disc disorders (40%) with the rest of the hospital controls.
Proxy interviewee for deceased cases, controls	0.7	0.4–1.0	Studies with population controls (18–22)	RRR for fatal outcomes relative to the RR for nonfatal outcomes from meta-regression, assuming an L-shaped dose-response curve.
Effect measure modification	0.6	0.4–0.8	Studies with population controls (18–22)	RRR for females relative to the RR for males from meta-regression.

* RRR, relative risk ratio, the ratio of the unadjusted relative risk (RR) to the adjusted relative risk; CI, confidence interval.

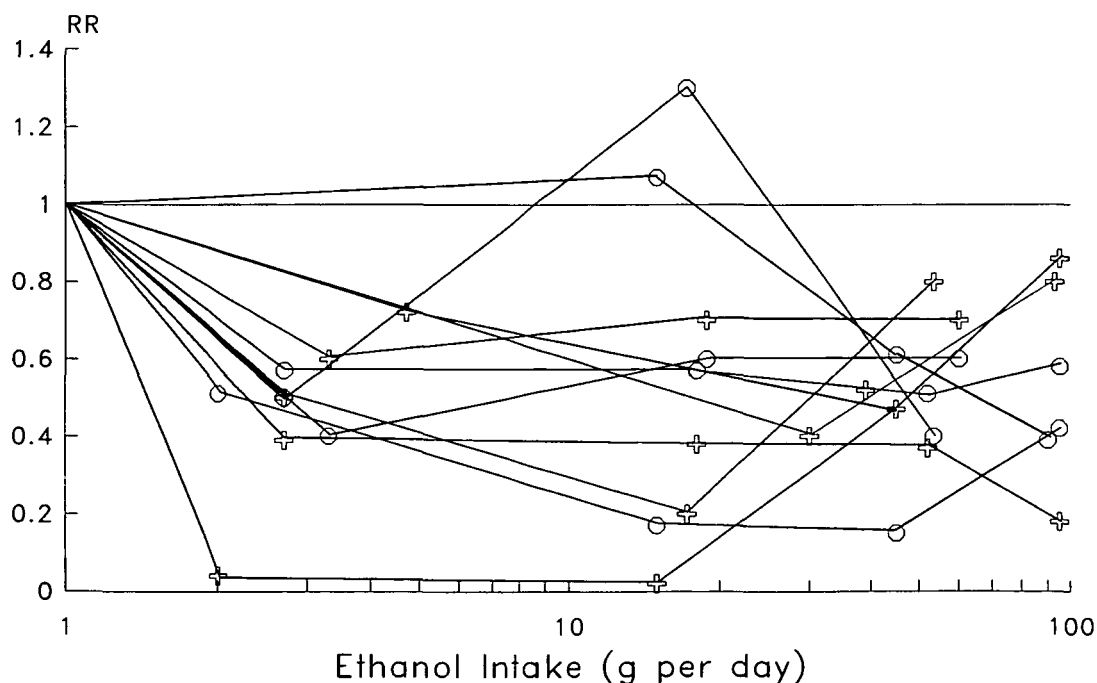


FIGURE 2. Estimates of the relative risk (RR) of coronary heart disease in relation to ethanol intake, from population-based case-control studies. Circles represent relative risks from studies of nonfatal myocardial infarction and crosses those from studies of fatal coronary heart disease.

nonfatal infarction ($\Delta = 1.10$; $p = 0.1$) and mixed outcomes ($\Delta = 1.15$; $p = 0.1$), but not well for fatal disease ($\Delta = 1.7$; $p < 0.001$). The fit for fatal outcomes was little improved ($\Delta = 1.7$) by inclusion of terms for

study flaws and characteristics (inclusion of ex-drinkers with nondrinkers, no control for age, no control for smoking, number of years of follow-up, number of exposure categories). Exclusion of the relative risks from

TABLE 3. Quadratic models of weighted logarithms of dose-specific relative risks from cohort studies of ethanol intake and risk of nonfatal myocardial infarction (MI), fatal coronary heart disease (CHD), and mixed CHD outcomes*

Outcome	Regression coefficients†				Goodness of fit (ρ)	Risk for ≥ 10 g/day (≥ 1 drink/day)‡	
	$\beta(X)$	SE§	$\beta(X^2)$	SE		RR§	95% CI§
Nonfatal	-21.3	3.5	0.157	0.042	0.1	0.82	0.76–0.89
Fatal CHD	-6.6	0.7	0.059	0.008	<0.001	0.94	0.91–0.97
Mixed CHD	-11.0	2.5	0.092	0.026	0.1	0.90	0.85–0.96

* Weights were inverses of covariance-adjusted variances of the logarithms.

† Change in the logarithm of the relative risk per g and per g² of ethanol intake.

‡ Reference group was nondrinkers.

§ SE, standard error ($\times 10^{-3}$); RR, relative risk; CI, confidence interval.

the American Cancer Society female cohort (32) reduced Δ to 1.5 ($p < 0.001$).

Incremental regression clarified the shape of the curve. Combining all cohort data, there appeared to be a decline in risk at doses up to one half drink per day, with little further change in risk associated with drinking more than one half drink per day (table 4 and figure 3). The resulting L-shape resembles a saturation effect, as if all of the hypothesized protective effect occurred at low doses.

Combining relative risks for all three types of outcomes, the fit of a quadratic model ($\Delta = 1.6$; $p < 0.001$) was no better than that for an L-shaped curve that was linear between 0 and 10 g of ethanol per day, with saturation (flattening) above 10 g/day ($\Delta = 1.6$). Addition of indicator terms for fatal and mixed outcomes, ex-drinkers in the nondrinker group, no control for age, no control for smoking, control for fewer than four variables, and female subjects reduced

TABLE 4. Incremental relative risks (iRRs) for each increment in ethanol intake, and corresponding dose-specific relative risks (RRs), from cohort studies of ethanol intake and coronary heart disease

Frequency of consumption (no. of drinks)	Ethanol intake (g/day)	iRR*	95% CI†	RR‡	Shifted 95% CI§
0	0			1.0	Reference
2/month	0.8	0.95	0.73–1.24	0.95	0.73–1.24
2/week	3.8	0.92	0.88–0.97	0.88	0.83–0.92
4/week	7.5	0.88	0.76–1.01	0.77	0.67–0.89
1/day	13	1.19	1.07–1.31	0.91	0.82–1.01
2–3/day	30	0.91	0.85–0.98	0.83	0.77–0.89
4–6/day	60	0.97	0.88–1.08	0.81	0.73–0.90
≥ 7 /day	120	1.17	1.02–1.34	0.95	0.83–1.09

* The iRR is the relative risk comparing consecutive levels of ethanol intake. The reference group is the subgroup of drinkers with less intake; the index group is the subgroup with greater intake. The iRRs were estimated by incremental weighted regression (a variation on moving line regression; see text) of logarithms of the RRs in the index interval, with weights equal to inverses of covariance-adjusted variances of each of the RRs in the interval. Instead of RR = 1 at 0 g/day, the origin for the regression was taken as the RR at a midpoint in the reference interval, i.e., at the level of intake specified in the first column of the preceding row of the table. The 95 percent confidence intervals were widened by multiplying the standard error by the residual deviance to account for heterogeneity among RRs in the index interval.

† CI, confidence interval.

‡ The RR for a given intake of ethanol was derived from the iRRs for the component increments (i.e., the product of the iRRs in the preceding rows).

§ The shifted 95 percent confidence interval is the 95 percent confidence interval of the iRR in the preceding row, shifted so that it is centered on the RR; it is a minimum estimate of the uncertainty of the RR.

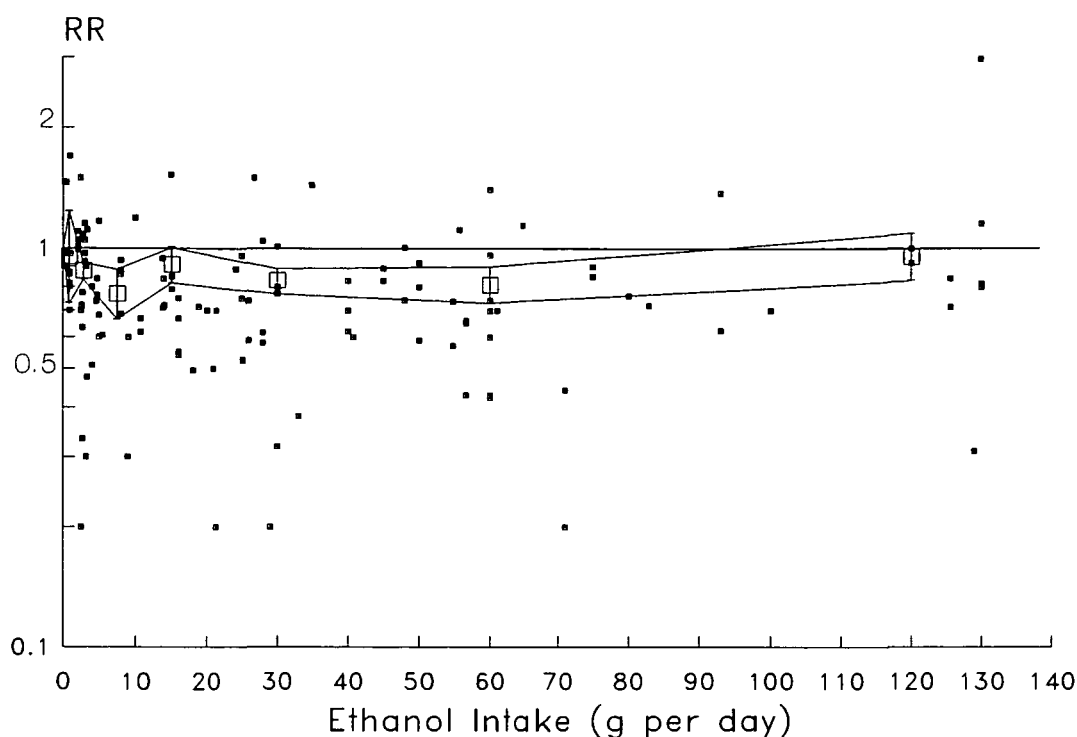


FIGURE 3. L-shaped relation between the relative risk (RR) of coronary heart disease and ethanol intake, based on meta-analysis of relative risks from all cohort studies (shown as points). Squares and error bars show RRs and 95% confidence intervals (adjusted for heterogeneity) of dose-specific weighted averages of RRs. The line was fitted by incremental weighted regression (a kind of moving line regression; see text), with weights equal to the inverse of the covariance-adjusted variance of the logarithm of the RR.

Δ to 1.4 ($p < 0.001$). After adjustment of the variances for this residual heterogeneity, the relative risk for intake of 6.5 g/day (a drink every other day) was 0.76 (95 percent confidence interval (CI) 0.65–0.90). This virtually refutes the null hypothesis that the association was “due to chance.” However, this means only that an unlikely scenario was refuted: that Nature had randomly assigned ethanol intake to the study subjects, yet there still occurred a chance imbalance of confounders (106).

To refute the possibility that the complexity of the meta-regression was obscuring the analysis, I also carried out simpler analyses. Weighted regression of low-dose data from five studies (23, 27, 35, 38, 42) that separated ex-drinkers from long-standing non-drinkers gave an estimated relative risk of 0.83 (95 percent CI 0.67–1.01) for one half drink per day in comparison with never

drinkers. Excluding nondrinkers altogether and using occasional drinkers as the reference group in five other studies (28, 29, 46–48), it was possible to estimate an incremental relative risk of 0.88 (95 percent CI 0.81–0.96) for approximately one drink per day relative to less-than-daily drinking: This agrees with the corresponding incremental relative risk of 0.91 (95 percent CI 0.88–0.95) from the American Cancer Society studies (31, 32).

These meta-regressions of cohort data confirm that a generalizable association *exists* (is not due to chance), but they do not explain *why* it exists.

Refutation of outcome misclassification bias

A recent review (107) of error in diagnosis of coronary heart disease concluded erroneously that cardiologic epidemiologic

studies are unreliable because their data are inaccurate. The author overlooked the distinction between insensitivity and nonspecificity of the diagnosis. By making this distinction, we shall see that outcome misclassification cannot account readily for the ethanol-infarction association.

Diagnosis of nonfatal myocardial infarction. It is estimated that half of all myocardial infarctions may be silent (108). In addition, many infarctions are fatal. Therefore, the diagnosis of nonfatal myocardial infarction is clearly very insensitive to the true incidence. However, a relative risk would be unaffected by insensitivity if 1) nonspecificity is negligible (109) and 2) the insensitivity is no greater among drinkers than among nondrinkers. The specificity of the diagnosis of myocardial infarction, when confirmed by chest pain, electrocardiography, and elevated cardiac enzymes, appears to be so high that leading research does not even discuss it (110). False positive diagnoses due to myocarditis are rare (111). On the other hand, insensitivity conceivably could be differential: The anesthetic effect of ethanol might mask infarction symptoms or influence survival. However, a reviewer (112) judged the evidence to be contradictory and inconclusive on the relation of ethanol intake to pain among angina patients. Moreover, differences in pain sensitivity would not readily explain the inverse association between ethanol intake and coronary death. Likewise, the inverse association with coronary death is opposite what one would expect if the association with nonfatal infarction was due to ethanol's causing poorer survival after infarction.

Diagnosis of fatal coronary disease. A diagnosis of coronary heart disease on a death certificate has less specificity than fatal acute myocardial infarction. For example, a major type of heart disease, sudden cardiac death, appears to have a positive predictive value of less than 80 percent as a proxy for fatal infarction, based on a study of resuscitated patients (113). Assuming that false positive diagnoses occur equally among nondrinkers and drinkers, calcula-

tions show that a relative risk of 0.8 for sudden fatal infarction among drinkers would be manifest as a relative risk of about 0.93 for sudden cardiac death as a whole. This would partly explain the weaker relation of ethanol to fatal outcomes than to nonfatal outcomes.

Loss to follow-up. Insensitivity of the outcome measure increases with loss to follow-up. However, in the major cohort studies (27–32), losses to follow-up were less than 3 percent. Only if loss were very strongly related to nonfatal disease would any bias be more than negligible in these studies.

In conclusion, we can rule out inaccuracy of diagnosis as an explanation for the apparent reduction in incidence of coronary disease at low ethanol intakes, but it could partly explain the flatness of the curve from moderate doses to high doses.

Refutation of exposure misclassification bias

Self-reported alcohol consumption is widely believed to be inaccurate. Differential recall, however, is mainly a problem for case-control studies, which we have already excluded. Exposure misclassification in prospective cohort studies would tend to be nondifferential (especially after stratification by risk group), because responses to the alcohol questions in the interview or on the questionnaire cannot be influenced by illness that has not yet occurred.

Underreporting. Most error in self-reported current alcohol intake is believed to be insensitivity, i.e., underreporting (114). This is sometimes compounded by insensitive questions. In the American Cancer Society study (31, 32) and the Multiple Risk Factor Intervention Trial (50), the "nondrinker" category must have included false negatives: drinkers who got their ethanol from beverages other than beer, wine, or whiskey. In the Puerto Rico Heart Health Program (26), the "unexposed" category included drinkers who happened to have abstained within 24 hours of the interview. Despite the resulting tendency for relative

risks to be biased toward the null, these studies showed significant associations.

Misclassification within drinkers. It is easier to remember whether or not you ever drank ethanol than to remember how often and how much. Therefore, incremental relative risks across different levels of drinking (excluding nondrinkers) might be more diluted toward the null, because of nondifferential misclassification, than relative risks for all drinkers versus nondrinkers. In the extreme, such unequal misclassification could cause various dose-response relations (continuously decreasing, U-shape) to be manifest as an L-shape.

How much flattening can be attributed to error in reporting of alcohol intake? Validation studies conducted within the Nurses' Health Study (115) and Health Professionals Follow-up Study (116) cohorts found that Spearman correlation coefficients of validity (comparing intakes measured by food frequency questionnaires and by the average of four 1-week diet records) were 0.90 in women and 0.86 in men. The amount of bias toward the null caused by nondifferential exposure misclassification is a function of the coefficient from a regression of "true" ethanol intake on the mismeasured variable (117). In the Nurses' Health Study, the coefficient was 0.67, which suggests that regression slopes may have underestimated the true slope by about 45 percent (117).

Irregular drinking. Most studies did not separate irregular drinkers from regular drinkers. From the British Regional Heart Study (48), we can calculate a relative risk of 1.2 (95 percent CI 0.83–1.9) among men who had 3–6 drinks per day on weekends only versus daily drinkers of 1–2 drinks per day. In the American Cancer Society study (31), irregular drinkers had an age- and smoking-adjusted relative risk of 1.2 (95 percent CI 1.1–1.3) relative to men who had 1–5 drinks per day. These data suggest that higher incidence among moderate irregular drinkers may obscure the apparent protective effect of regular drinking.

Drift. Although underestimation (insensitivity) is probably the main type of error in

measuring current drinking, overestimation (nonspecificity) may be a problem when current drinking is used as a proxy for future drinking. Mean intake of alcohol declines with age in cross-sectional data, but in the United States, consumption per capita increased in the 1960s and 1970s and declined in the 1980s (118). Intake increased in the Framingham Study cohort even among patients with prior diagnoses of heart disease (46). In the Alameda County Study (42), drift of people away from their baseline status as a drinker or nondrinker was equal to a decline in positive predictive value of about 1 percent per year of follow-up, and a decline in negative predictive value of about 3 percent per year. Over 10 years of follow-up, this would produce a 20 percent dilution corresponding to a relative risk ratio of about 1.08. However, in the more recent large cohort studies, durations between exposure assessment and outcome were too short for drift to be important (Health Professionals Follow-up Study, 1.5 years; Nurses' Health Study, 4 years; Kaiser Permanente cohort study, 2.5 years).

Ex-drinkers. Shaper (119) hypothesized that the U-shaped mortality curve was due to contamination of the nondrinkers with ex-drinkers. Our meta-analysis allays concern. The weighted average of the relative risks for ex-drinkers relative to nondrinkers in five cohort studies (23, 28, 38, 42, 77) plus results from the Health Professionals Study (Eric Rimm, Harvard University, personal communication, 1992) was 1.07 (95 percent CI 0.89–1.28). Therefore, although it is desirable to separate ex-drinkers from long-term abstainers, failure to do so would usually cause minimal bias. In the meta-regression, the coefficient for failure to separate ex-drinkers corresponded to a relative risk ratio of 1.05 (table 5). Moreover, both the Nurses' Health Study (28) and the Health Professionals Study (29) reported that the relative risks did not change substantially after exclusion of subjects who reported recent changes in their alcohol intake.

TABLE 5. Estimated relative risk ratios for hypothesized sources of bias and confounding in cohort studies of ethanol intake and risk of coronary heart disease (CHD)

Source	RRR*	95% CI*	Studies	Estimation
Publication bias	1.02		Coffee and CHD (91–97)	Assuming a weakly positive association (RR = 1.2) in seven cohort studies not included.
Nonspecificity of sudden cardiac death	1.17		Resuscitation study (113)	Based on an estimated positive predictive value of 0.8.
Nonspecificity			All cohort studies	Meta-regression coefficient:
Fatal CHD	1.4	1.2–1.6		RR for fatal/RR for nonfatal
Mixed CHD	1.2	1.0–1.4		RR for mixed/RR for nonfatal
Nondifferential error in self-reports of ethanol intake	1.17		Nurses' Health Study (28), Health Professionals Study (29)	Based on correlation coefficients of validity of 0.9 for ethanol use; questionnaire vs. four 7-day diet records.
Drift from exposure level at baseline	1.07		Alameda County Study (42)	Positive predictive value dropped 1%/year; negative predictive value, 3%/year.
Confounding by age	0.67		Japanese physicians (65)	Crude RR divided by age-adjusted RR.
	0.97			Residual confounding due to use of 15-year age categories.
	0.62		Kaiser Permanente cohort (27)	Crude RRs/age-adjusted RRs; for rare drinkers, occasional drinkers, and ≥ 1 drink/day drinkers.
	0.44			
	0.56		All cohort studies	RRs from studies without age adjustment, compared with age-adjusted RRs by meta-regression.
	0.92	0.5–1.8		
Confounding by smoking	1.09		American Cancer Society (31)	Age-adjusted RR/age- and smoking-adjusted RR (1–2 drinks/day).
	1.0		Japanese physicians (65)	As above (< 2 drinks/day)
	1.17			As above (≥ 2 drinks/day)
	1.02		Framingham Study (120)	Residual confounding due to smoker misclassification due to drift from baseline.
	1.1	0.79–1.6	All cohort studies	RRs only age-adjusted, compared with age- and smoking-adjusted RRs by meta-regression assuming an L-shape.
Confounding by age, smoking, and socioeconomic status combined	1.04–1.09		American Cancer Society (31)	Crude RRs/age-, smoking-, and education-adjusted RRs.
	0.95–1.0		British Regional Heart Study (48)	Crude RRs/age-, smoking-, and socioeconomic status-adjusted RRs.
Confounding by body mass index	0.85		Kaiser Permanente Study (77)	Change in RR on adjustment for body mass index.
Confounding by history of chronic illness	0.9; 0.9; 0.95–1.1		Alameda County (42); Italian rural cohorts (81); Health Professionals Study (29)	Based on change in RR after exclusion of people with chronic illnesses from the cohorts at baseline.
Contamination of nondrinkers by ex-drinkers	1.0		Kaiser Permanente Study (77)	RR = 1.0 for ex-drinkers, after adjustment for age, sex, race, smoking, body mass index, education, and marital status.
	1.05	1.0–1.1	All cohort studies	Antilog of meta-regression coefficient for RRs from studies with ex-drinkers included among nondrinkers.

* RRR, relative risk ratio, the ratio of the unadjusted relative risk (RR) to the adjusted relative risk; CI, confidence interval (where calculable).

Refutation of hypothesized confounders

In this section, we attempt to quantify uncontrolled confounding by known risk factors. Unfortunately, only a few studies provide data that enable estimation of relative risk ratios due to confounding. Moreover, the generalizability of such ratios is probably lower than that of relative risks, because they depend on the magnitudes of culturally determined associations between alcohol intake and coronary risk factors.

Age. How much of the L-shape from our meta-regression could be due to residual confounding by age? In the study of Japanese physicians (65), 15-year age categories were used. If 5-year age categories had been used, the fully adjusted relative risk would have been about 0.91 instead of 0.88, which gives a relative risk ratio of 0.97 for residual confounding by age. Comparing this to the relative risk ratio of 0.67 (the crude relative risk divided by the age-adjusted relative risk) shows that only about 7 percent of the original confounding by age remained uncontrolled. Therefore, since most studies used narrower age categories than the Japanese study, residual confounding by age seems very unlikely to be an explanation for the L-shape.

Smoking. Adjusting for smoking causes the ethanol-infarction association to be stronger, not weaker. The Japanese physician study (65) gave both age-adjusted and age- and smoking-adjusted relative risks for fatal coronary disease. This permits calculation of a relative risk ratio of 1.2 due to smoking by heavier drinkers. Among lighter drinkers, the ratio was 1.0, meaning no confounding. In the American Cancer Society study (31), the ratio among men who reported having one or two drinks per day was 1.1.

How much of the flattening of the curve at higher doses of ethanol might be due to residual confounding by smoking? Some residual confounding would result from misclassification of ex-smokers as smokers. In the Framingham Study (120), it was shown that 40 percent of nondrinkers who had been

smokers at baseline were ex-smokers after 10 years of follow-up, whereas among drinkers the figure was 30 percent. Calculations suggest that this would result in a relative risk ratio of only 1.02 due to residual confounding by smoking. Additional residual confounding from other kinds of error in quantifying smoking would tend to bias the relative risk further upward.

Education and socioeconomic status. Two reports (31, 48) enabled calculation of relative risk ratios of 0.95–1.09 for net confounding by age, smoking, and education or socioeconomic status, but ratios for education or socioeconomic status alone were not calculable (table 5). It appears unlikely that residual confounding by education or socioeconomic status is an explanation for the association, but it is possible that these variables are poor proxies for an unmeasured confounder with a high relative risk ratio.

Obesity. If obese people were more likely to be nondrinkers, the ethanol-infarction association could be due partly to residual confounding by obesity. A review (101) of 51 studies of the relation between alcohol intake and adiposity showed no clear overall pattern. The Kaiser Permanente study (77), the Honolulu Heart Study (121), and the Nurses' Health Study (122) found that nondrinkers were more likely to be obese, but this was not found in the Health Professionals Study (122). In the Kaiser Permanente study (77), adjustment for body mass index caused the relative risk for fatal heart disease to attenuate from 0.56 to 0.66, which gives a relative risk ratio of 0.85. Weight loss was the reason given by 15 percent of people who reported having reduced their intake of alcohol. This subgroup had a relative risk of 1.4 (95 percent CI 0.97–2.0), adjusted for age, sex, race, smoking, body mass index, marital status, and education.

Physical activity. A more likely source of residual confounding is physical activity, because it is much more difficult to measure than obesity. In fact, only four cohort studies (26, 28, 42, 68) adjusted for measures of physical activity. In the Alameda County

Study (42), the adjustment had a negligible effect on the relative risk, but this could be merely because the index of physical activity was poor. A meta-analysis (100) concluded that people in sedentary occupations have a relative risk of 1.9 (95 percent CI 1.6–2.2) for coronary disease when compared with people in active occupations. This is comparable to the relative risk for smoking. If the association between a sedentary lifestyle and nondrinking were as strong as that between nonsmoking and nondrinking, the relative risk ratio for a sedentary lifestyle would be between 1.0 and 1.1.

Diet. The vast literature on diet and heart disease suggests that multiple nutrients influence risk (123). The Honolulu (23), Nurses' (28), and Health Professionals (29) studies adjusted for multiple dietary risk factors, including saturated and polyunsaturated fats and cholesterol. Adjustment for nutrients had almost no effect. This is consistent with the lack of strong associations between ethanol and nutrient intake in the Honolulu (121), Nurses' (122), and Health Professionals (122) cohorts.

The latest hypothesis to be corroborated in more than one study is that antioxidants, such as vitamin E and beta-carotene, may be protective. In the Health Professionals cohort (124), ethanol intake was only slightly higher in the lowest quintile of vitamin E intake. In the Nurses' Health Study (125), there was no association between intakes of ethanol and vitamin E.

The meta-analysis (88) of coffee intake and risks of fatal and nonfatal heart disease found that the relative risks tended to be elevated among drinkers of five or more cups per day. In some studies, alcohol drinkers consume more coffee than nondrinkers (126). Like smoking, this would tend to flatten the curve at higher doses but would not explain the reduction in risk at low doses of ethanol. Two studies (23, 27) adjusted for coffee intake and found that it did not affect the association with ethanol.

Medications. Confounding by medications is substantially avoided by restricting cohorts to subjects without clinical disease.

However, in a healthy population, aspirin use (127) could magnify the ethanol-infarction association if nondrinkers abstained from aspirin while occasional drinkers took it frequently. Current use of postmenopausal estrogens in the Nurses' Health Study (128) was associated with an average ethanol intake of 7.9 g/day, compared with 7.5 g/day among former users and 7.3 g/day among never users. This would slightly exaggerate the ethanol-infarction association among women because of the cardioprotective effect of estrogen (128).

Intermediate biochemical markers. A physiologic coronary risk factor such as high density lipoprotein cholesterol, blood pressure, or diabetes can be both a confounder and an intermediate in the causal pathway between ethanol intake and infarction. Not adjusting for it would leave residual confounding; adjusting for it would tend to underestimate the association (129). Special methods to control for confounding by intermediates have only recently been developed (129) and were not used in any of the studies reviewed here. Criqui et al. (39) reported that adjustment for high density lipoprotein cholesterol reduced but did not eliminate the ethanol-infarction association. In the Kaiser Permanente cohort study (27), relative risk ratios for blood pressure, serum cholesterol, and glucose ranged from 0.98 to 0.90 when these variables were added to models that had already adjusted for age, smoking, sex, race, coffee intake, and education. The Nurses' Health Study (130) found that ethanol may be protective against the original development of diabetes, which would mean that adjusting for diabetes, as in several cohort studies (28, 29, 43, 44, 50, 52, 53), would result in some underestimation of the effect of ethanol.

Refutation of reverse causation

Shaper (119) hypothesized that existing diseases cause drinkers to quit, such that the nondrinking category becomes contaminated by ill people. If the illness is heart disease itself, this would be reverse causation bias. Indeed, in the Kaiser Permanente

study (77), about 40 percent of ex-drinkers said they had quit for medical reasons. Moreover, data from the Alameda County Study (42) and the Italian part of the Seven Countries Study (81) indicate relative risk ratios of 0.9 if history of chronic illness is treated as a confounder. However, Shaper's hypothesis was refuted by the large cohort studies (28, 29, 31, 77), which found that the ethanol-infarction association persisted when analyses excluded ex-drinkers and/or people with chronic illnesses. In the American Cancer Society study (31), among the 33 percent of men who were "sick at enrollment," the prevalence of drinking was actually the same as in the rest of the cohort, and the relative risks for coronary death were virtually identical. In the Health Professionals Study (29), which excluded men entirely if they had a history of cancer, myocardial infarction, angina, stroke, or cardiac procedures, an additional analysis was carried out after further exclusions: history of gout, diabetes, hypercholesterolemia, hypertriglyceridemia, hypertension, or other heart problems. The resulting relative risk ratios were 1.1 among occasional drinkers and 0.95 among light drinkers.

In the Kaiser Permanente cohort (77), the 50 percent excess in age-adjusted incidence of fatal cardiovascular outcomes among ex-drinkers as compared with never drinkers disappeared after adjustment for sex, race, smoking, body mass index, marital status, and education. Similar adjustments did not eliminate the excess incidence of noncardiovascular death (relative risk = 1.3, 95 percent CI 1.0–1.7). This suggests that Shaper's concerns may be valid for studies of total mortality but not for coronary mortality.

Reverse causation could have biased the association in several case-control studies (13, 19–21) that included patients who had had previous diagnoses of myocardial infarction.

Refutation of unmeasurable confounders and biases

In addition to the confounders and biases mentioned above, there are other potential explanatory factors that were not measured

or controlled: stress, personality, lying, physical activity at work, and nonethanol ingredients of alcoholic drinks. Hill's (131) criteria for causal inference are useful as weak tests of these alternatives. Although they are commonly used as a checklist of *supporting* evidence, we shall see how the criteria are better treated as *criteria for refuting* unmeasurable or unknown confounding and bias.

Strength of association. A strong association is harder to explain away as being due to an unnoticed alternative risk factor. By Hill's definition of strong (131), the relation in the ethanol-infarction association is weak. Therefore, it is reasonable to hypothesize that life events, Type A behavior, and anger, which are hypothesized to be risk factors for heart disease (132), might be confounders. These phenomena are difficult to measure and, indeed, none of the cohort studies mentioned controlling for them. Evidence for only weak confounding comes from a meta-analysis of Type A behavior and risk of coronary disease which concluded that the relation was found only in a subset of studies and was weakened if the meta-analysis weighted studies by size (102). There is some evidence that Type A men drink more alcohol than Type B men (103). Type A behavior would have to be very strongly associated with ethanol intake to produce substantial confounding.

Consistency. Several hypotheses concerning unmeasured confounders are refuted by the fact that the ethanol-infarction relation is seen consistently across diverse populations. If diversity is defined as "large between-population variability in the magnitude and direction of associations between ethanol intake and confounders," then diversity can be viewed as quasi-randomization by Nature. This author believes that diversity tends to increase the variability of the magnitude of confounding across studies. There is no reason why the confounding would tend to cancel out with large numbers of studies (as it would in a randomized trial), but if the magnitude of the ethanol-infarction association is consistent across diverse populations, it is harder

to explain it away as being due to unmeasured confounders.

The hypothesis that other ingredients of certain alcoholic beverages are the actual preventives, not the ethanol itself, is refuted by the consistency of the association across different types of alcoholic drinks. Weighted averaging of the relative risks reported for beer, wine, and liquor separately in five cohort studies (23, 28, 36, 46, 77), plus data from the Health Professionals Study (Eric Rimm, Harvard University, personal communication, 1992), gave almost identical relative risks: 0.78 (95 percent CI 0.70–0.87) for beer, 0.74 (95 percent CI 0.65–0.85) for wine, and 0.79 (95 percent CI 0.72–0.86) for liquor.

Cross-cultural consistency helps rule out distortion due to drinkers' lying about their intake and confounding due to an undiscovered dietary factor or health-related behavior. Reasons for not drinking would probably be quite different among elderly Americans who lived through the Prohibition era than among middle-aged Japanese physicians. The more diverse the cultures are, the more difficult it is to explain the association as being due to a cultural allocation bias.

The association was consistent between the sexes. A weighted average of relative risks from studies that included women (27, 28, 35, 37, 39, 41–44, 46, 53, 58), excluding the American Cancer Society study, yielded a relative risk of 0.81 (95 percent CI 0.66–0.98) for occasional drinkers versus nondrinkers. In the female portion of the American Cancer Society study (32), the corresponding relative risk was 0.87 (95 percent CI 0.80–0.93). Regression estimates of the relative risk ratio for females versus males ranged from 0.96 to 1.16. This suggests that substantial confounding by sex-specific exposures such as estrogen use did not occur.

Consistency across the age spectrum was also seen. A weighted average of the relative risks for people over 65 from six studies (35, 37, 57, 52, 77, 133) yielded a relative risk of 0.78 (95 percent CI 0.72–0.86) for drinkers

versus nondrinkers. This, plus the consistency of the association in the 1960s, 1970s, and 1980s and in both sick and healthy populations, helps further rule out confounding by physical activity, medications, or illnesses.

Dose response. The existence of a monotonic trend in risk, i.e., a continuously increasing or decreasing dose-response relation (or "biologic gradient"), is a special case of the criterion of consistency across diverse subgroups. Here groups differ in their doses of ethanol. The monotonic trend at low doses (table 4) from incremental regression refutes the hypothesis that the ethanol-infarction association was entirely due to contamination of the nondrinking group by ex-drinkers.

If the monotonic trend had continued beyond the low-dose range, many of the hypothesized biases and uncontrolled confounders would have been further discredited. As it is, the L-shape leaves critics a toehold. It is still possible to argue that there is confounding by some characteristic of people who seldom or never drink alcohol.

Plausible mechanism. The lack of a smooth dose-response relation may mean that some confounders are more difficult to refute, but it does not rule out causation. The relation between aspirin intake and coronary risk is believed to be L-shaped, and a plausible mechanism has been demonstrated: Aspirin taken at low doses every other day produces a sustained reduction in platelet aggregability—i.e., a saturation curve (127). It is plausible that ethanol at low doses has a similar effect (134).

A refutationist is suspicious of mechanistic arguments used to bolster the hypothesis of interest. Literature on the biomedical effects of ethanol illustrates the need for suspicion (135). Ethanol appears to have many effects on the circulatory system, some beneficial, some adverse. A mechanism can be found to support any prejudice.

The main use of the plausibility criterion should be to cast doubt on mechanisms that

are implausible because of their complexity. An example of a less plausible mechanism is the hypothesis that psychological characteristics of nondrinkers and heavy drinkers predispose them to stress, and consequently heart disease, whereas occasional and moderate drinkers are more emotionally adjusted. This hypothesis is convoluted with multiple steps, each of which provides many opportunities for the association to be diluted by other causes. By comparison, the direct biochemical effects of ethanol are more plausible explanations, because they involve fewer intermediate steps where other causal factors might intervene.

Coherence. Plausibility is reduced by "incoherent" evidence. For example, the ethanol-infarction hypothesis was once criticized because evidence seemed to suggest that subfraction 2 of high density lipoprotein cholesterol (HDL₂) was more protective than subfraction 3 (HDL₃). This was incoherent with the observation that ethanol elevated HDL₃ more than HDL₂. The incoherence has since been resolved. Current evidence suggests that HDL₃ is at least as protective as HDL₂ (136).

Analogy. Arguments for mechanistic plausibility often draw on analogies. For example, ethanol appears to have an analogous, apparently protective effect on risk of symptomatic gallstones (104), a disease which shares other etiologic factors with coronary heart disease, such as obesity, low vegetable intake, and possibly smoking. Like the criterion of plausibility, analogy is easy to misuse. The use of analogy in a refutationist analysis is mainly for generating competing hypotheses. For example, our concern about selection bias in case-control studies was based largely on analogy with other case-control studies in which selection bias has been documented (137).

Specificity. Ethanol lacks "specificity of effect," because it has multiple biologic and behavioral effects. Consequently, it is easy to hypothesize biases in case-control studies (e.g., selection bias and recall bias) and in cohort studies (e.g., differential loss to follow-up and residual confounding). In ad-

dition, coronary heart disease lacks "specificity of cause," because it has a multifactorial etiology. This means we can easily add plausible hypotheses to the list of potential confounders. Specificity is a somewhat tautologic criterion: It amounts to a restatement of the principle that causal inference is contingent on lack of alternative explanations.

Temporality. The principle that cause must precede effect is used to refute reverse causation. Earlier in this review, we saw that reverse causation was refuted by restricting cohorts to subjects who reported no history of chronic disease.

DISCUSSION

Deductive meta-analysis of evidence for and against more than 20 hypotheses concerning the relation between ethanol intake and incidence of myocardial infarction corroborated the preventive hypothesis by weakening competing hypotheses. The decline in risk at low doses does not appear to this author to be due to random sampling variation, selection bias, reverse causation, or error in measuring ethanol intake or heart disease. Confounding by an unidentified risk factor that is common in nondrinkers but not in occasional drinkers is hard to imagine. Residual confounding by a combination of factors—obesity, sedentary lifestyle, aspirin, and diet—is difficult to rule out, but when I construct such a mixed hypothesis, it seems too contrived.

As for the flattening of risk with ethanol intake greater than one drink every other day, several competing explanations remain. One is that multiple effects of ethanol on blood cancel each other out and the overall result is a "saturation effect," a true flattening of the preventive relation. Another is that error in measuring ethanol intake and confounders among heavier drinkers causes a spurious flattening of the curve. Therefore, the U-hypothesis is not yet conclusively refuted.

Clinical advice

This analysis suggests that people who have 2–4 alcoholic drinks per day can safely cut their intake to one drink per day. On the other hand, most health professionals still refrain from suggesting that nondrinkers start drinking small quantities of ethanol (138). Equally effective prevention of heart disease may be achieved by other means (139), including control of weight and blood pressure, use of aspirin and possibly antioxidant supplements, moderate exercise, and intake of oleic acid instead of saturated and trans fatty acids.

Meta-analysis controversy

This review has demonstrated that a refutationist approach to epidemiologic inference is a solution to the problem of overinterpretation of meta-analysis in epidemiology. Meta-analyses should be designed as tests of competing explanations, not mere summaries of summaries. Meta-analyses of randomized double-blind trials have the luxury of focusing on the refutation of chance (random imbalance of net confounding), because selection bias, information bias, and nonchance confounding are minimized by randomization and blinding. Meta-analyses of nonexperimental studies are more difficult, because they must refute many more competing hypotheses before causation can be inferred.

The controversy (1, 140–143) about meta-analysis of nonexperimental studies is in large part a reaction to the overinterpretation of confidence intervals that exclude the null value and overreliance on the consistency criterion. The narrowness of meta-confidence intervals merely forces us to grapple with the fact that confidence intervals in nonexperimental studies represent only one type of uncertainty, the meaning of which is obscure (106). The consistency criterion corroborates the causal hypothesis of interest only indirectly by refuting confounders and biases that differ across studies. It does not refute confounders and biases that recur consistently in many studies.

Remaining problems

This meta-analysis could have been more rigorously deductive. Many competing hypotheses, particularly that of publication bias, were treated only semiquantitatively. A more thorough search of the literature for ancillary data would have been desirable given additional resources. With better estimates of relative risk ratios (including their variances and heterogeneity across studies), it would have been fruitful to perform sensitivity analyses of the effect of dividing observed relative risks by relative risk ratios to adjust for hypothesized study flaws.

Incompleteness is an inevitable characteristic of a deductive meta-analysis, for the same reasons that the combination of selected variables, transformations, and modeling assumptions is virtually limitless in multivariate analysis (89). The open-endedness reflects the manner by which generalizable knowledge grows, but poses a challenge for authors' time, editors' space, and readers' interest. A quick and simple meta-synthesis is more intelligible at the risk of being misleading. A long and complex deductive approach is more rigorous at the risk of being unintelligible. The optimum combination of parsimony and rigor will depend on the number of competing hypotheses, the quantity of evidence available, and social costs of the policy alternatives.

ACKNOWLEDGMENTS

This work was supported by a grant from Anheuser-Busch, Inc. (St. Louis, MO), administered through the Washington Technical Information Group, Inc., Washington, DC.

The author thanks Drs. Sander Greenland and Matthew Longnecker for their helpful suggestions.

REFERENCES

1. Dickersin K, Berlin JA. Meta-analysis: state-of-the science. *Epidemiol Rev* 1992;14:154–76.
2. Maclure M. Multivariate refutation of aetiological hypotheses in non-experimental epidemiol-

- ogy. *Int J Epidemiol* 1990;19:782-7.
3. Kneale W, Kneale M. The development of logic. Oxford, England: Clarendon Press, 1962: 36,175.
 4. Buck C. Popper's philosophy for epidemiologists. *Int J Epidemiol* 1975;4:159-68.
 5. Maclure M. Popperian refutation in epidemiology. *Am J Epidemiol* 1985;121:343-50.
 6. Weed DL. On the logic of causal inference. *Am J Epidemiol* 1986;123:965-79.
 7. Rothman KJ, Lanes SF, eds. Causal inference. Chestnut Hill, MA: Epidemiology Resources, 1988.
 8. Pearce NE, Crawford-Brown DJ. Critical discussion in epidemiology: problems with the Popperian approach. *J Clin Epidemiol* 1989;42: 177-84.
 9. Ng SK. Does epidemiology need a new philosophy? A case study of logical inquiry in the acquired immunodeficiency syndrome epidemic. *Am J Epidemiol* 1991;133:1073-7.
 10. Greenland S. Quantitative methods in the review of epidemiologic literature. *Epidemiol Rev* 1987;9:1-30.
 11. Marmot M, Brunner E. Alcohol and cardiovascular disease: the status of the U shaped curve. *BMJ* 1991;303:565-8.
 12. Moore RD, Pearson TA. Moderate alcohol consumption and coronary artery disease. *Medicine (Baltimore)* 1986;65:242-67.
 13. Stason WB, Neff RK, Miettinen OS, et al. Alcohol consumption and nonfatal myocardial infarction. *Am J Epidemiol* 1976;104:603-8.
 14. Rosenberg L, Slone D, Shapiro S, et al. Alcoholic beverages and myocardial infarction in young women. *Am J Public Health* 1981;71: 82-5.
 15. Kaufman DW, Rosenberg L, Helmrich SP, et al. Alcoholic beverages and myocardial infarction in young men. *Am J Epidemiol* 1985;121: 548-54.
 16. La Vecchia C, Franceschi S, Decarli A, et al. Risk factors for myocardial infarction in young women. *Am J Epidemiol* 1987;125:832-43.
 17. Kalandidi A, Tzonou A, Toupadaki N, et al. A case-control study of coronary heart disease in Athens, Greece. *Int J Epidemiol* 1992;21: 1074-80.
 18. Kono S, Handa K, Kawano T, et al. Alcohol intake and nonfatal acute myocardial infarction in Japan. *Am J Cardiol* 1991;68:1011-14.
 19. Scragg R, Stewart A, Jackson R, et al. Alcohol and exercise in myocardial infarction and sudden coronary death in men and women. *Am J Epidemiol* 1987;126:77-85.
 20. Jackson R, Scragg R, Beaglehole R. Alcohol consumption and risk of coronary heart disease. *BMJ* 1991;303:211-16.
 21. Hennekens CH, Rosner B, Cole DS. Daily alcohol consumption and fatal coronary heart disease. *Am J Epidemiol* 1978;107:196-200.
 22. Siscovick DS, Weiss NS, Fox N. Moderate alcohol consumption and primary cardiac arrest. *Am J Epidemiol* 1986;123:499-503.
 23. Yano K, Rhoads GG, Kagan A. Coffee, alcohol and risk of coronary heart disease among Japanese men living in Hawaii. *N Engl J Med* 1977; 297:405-9.
 24. Kagan A, Yano K, Rhoads GG, et al. Alcohol and cardiovascular disease: the Hawaiian experience. *Circulation* 1981;64(suppl 3):27-31.
 25. Kozarevic D, Demirovic J, Gordon T, et al. Drinking habits and coronary heart disease: The Yugoslavia Cardiovascular Disease Study. *Am J Epidemiol* 1982;116:748-58.
 26. Kittner SJ, Garcia-Palmieri MR, Costas R Jr, et al. Alcohol and coronary heart disease in Puerto Rico. *Am J Epidemiol* 1983;117:538-50.
 27. Klatsky AL, Armstrong MA, Friedman GD. Relations of alcoholic beverage use to subsequent coronary artery disease hospitalization. *Am J Cardiol* 1986;58:710-14.
 28. Stampfer MJ, Colditz GA, Willett WC, et al. A prospective study of moderate alcohol consumption and the risk of coronary disease and stroke in women. *N Engl J Med* 1988;319: 267-73.
 29. Rimm EB, Giovannucci EL, Willett WC, et al. Prospective study of alcohol consumption and risk of coronary disease in men. *Lancet* 1991; 338:464-8.
 30. Kivela S-L, Nissinen A, Ketola A, et al. Alcohol consumption and mortality in aging or aged Finnish men. *J Clin Epidemiol* 1989;42:61-8.
 31. Boffetta P, Garfinkel L. Alcohol drinking and mortality among men enrolled in an American Cancer Society prospective study. *Epidemiology* 1990;1:342-8.
 32. Garfinkel L, Boffetta P, Stellman SD. Alcohol and breast cancer: a cohort study. *Prev Med* 1988;17:686-93.
 33. Dyer AR, Stamler J, Paul O, et al. Alcohol consumption and 17-year mortality in the Chicago Western Electric Company Study. *Prev Med* 1980;9:78-90.
 34. Kozararevic D, McGee D, Vojvodic N, et al. Frequency of alcohol consumption and morbidity and mortality: The Yugoslavia Cardiovascular Disease Study. *Lancet* 1980;1:613-16.
 35. Klatsky AL, Friedman GD, Siegelau AB. Alcohol and mortality: a ten-year Kaiser-Permanente experience. *Ann Intern Med* 1981;95: 139-45.
 36. Salonen JT, Puska P, Nissinen A. Intake of spirits and beer and risk of myocardial infarction and death—a longitudinal study in eastern Finland. *J Chronic Dis* 1983;36:533-43.
 37. Colditz GA, Branch LG, Lipnick RJ, et al. Moderate alcohol and decreased cardiovascular mortality in an elderly cohort. *Am Heart J* 1985; 109:886-9.
 38. Kono S, Ikeda M, Tokudome S, et al. Alcohol and mortality: a cohort study of male Japanese physicians. *Int J Epidemiol* 1986;15:527-32.
 39. Criqui MH, Cowan LD, Tyroler HA, et al. Lipoproteins as mediators for the effects of alcohol consumption and cigarette smoking on cardiovascular mortality: results from the Lipid Research Clinics Follow-up Study. *Am J Epidemiol* 1987;126:629-37.
 40. Suhonen O, Aromaa A, Reunanen A, et al. Alcohol consumption and sudden coronary death in middle-aged Finnish men. *Acta Med Scand*

- 1987;221:335-41.
41. Camacho TC, Kaplan GA, Cohen RD. Alcohol consumption and mortality in Alameda County. *J Chronic Dis* 1987;40:229-36.
 42. Lazarus NB, Kaplan GA, Cohen RD, et al. Change in alcohol consumption and risk of death from all causes and from ischaemic heart disease. *BMJ* 1991;303:553-6.
 43. Leaverton PE, Havlik RJ, Ingster-Moore LM, et al. Coronary heart disease and hypertension. In: Cornoni-Huntley JC, Huntley RR, Feldman JJ, eds. *Health status and well-being of the elderly: National Health and Nutrition Examination Survey I Epidemiologic Follow-up Study*. New York: Oxford University Press, 1990:53-70.
 44. Semenciw RM, Morrison HI, Mao Y, et al. Major risk factors for cardiovascular disease mortality in adults: results from the Nutrition Canada Survey cohort. *Int J Epidemiol* 1988;17:317-24.
 45. Marmot MG, Rose G, Shipley MJ, et al. Alcohol and mortality: a U-shaped curve. *Lancet* 1981;1:580-3.
 46. Gordon T, Kannel WB. Drinking habits and cardiovascular disease: The Framingham Study. *Am Heart J* 1983;105:667-73.
 47. Gordon T, Doyle JT. Drinking and mortality: The Albany Study. *Am J Epidemiol* 1987;125:263-70.
 48. Shaper AG, Phillips AN, Pocock SJ, et al. Alcohol and ischaemic heart disease in middle aged British men. *Br Med J* 1987;294:733-7.
 49. Miller GJ, Beckles GLA, Maude GH, et al. Alcohol consumption: protection against coronary heart disease and risks to health. *Int J Epidemiol* 1990;19:923-30.
 50. Shaten BJ, Kuller LH, Neaton JD. Association between baseline risk factors, cigarette smoking and CHD mortality after 10.5 years. *Prev Med* 1991;20:655-9.
 51. De Labry LO, Glynn RJ, Levenson MR, et al. Alcohol consumption and mortality in an American male population: recovering the U-shaped curve—findings from the Normative Aging Study. *J Stud Alcohol* 1992;53:25-32.
 52. Scherr PA, LaCroix AZ, Wallace RB, et al. Light to moderate alcohol consumption and mortality in the elderly. *J Am Geriatr Soc* 1992;40:651-7.
 53. Cullen KJ, Knuiam MW, Ward NJ. Alcohol and mortality in Busselton, Western Australia. *Am J Epidemiol* 1993;137:242-8.
 54. Hein HO, Sorensen H, Suadicani P, et al. Alcohol consumption, Lewis phenotypes, and risk of ischaemic heart disease. *Lancet* 1993;341:392-6.
 55. Keller M, ed. *Second special report to the US Congress on alcohol and health: new knowledge*. Washington, DC: US GPO, 1974: 68-92. (DHEW publication no. (ADM) 75-212).
 56. Klatsky AL, Friedman GD, Siegelau AB. Alcohol consumption before myocardial infarction: results from the Kaiser-Permanente epidemiologic study of myocardial infarction. *Ann Intern Med* 1974;81:294-301.
 57. Dyer AR, Stamler J, Paul O, et al. Alcohol consumption, cardiovascular risk factors, and mortality in two Chicago epidemiologic studies. *Circulation* 1977;56:1067-74.
 58. Petitti DB, Wingerd J, Pellegrin F, et al. Risk of vascular disease in women: smoking, oral contraceptives, noncontraceptive estrogens, and other factors. *JAMA* 1979;242:1150-4.
 59. Klatsky AL, Friedman GD, Siegelau AB. Alcohol use, myocardial infarction, sudden cardiac death, and hypertension. *Alcohol Clin Exp Res* 1979;3:33-9.
 60. Thomas CB, Santora PB, Shaffer JW. Health of physicians in midlife in relation to use of alcohol: a prospective study of a cohort of former medical students. *Johns Hopkins Med J* 1980;146:1-10.
 61. Greig M, Pemberton J, Hay I, et al. A prospective study of the development of coronary heart disease in a group of 1202 middle-aged men. *J Epidemiol Community Health* 1980;34:23-30.
 62. Dyer AR, Stamler J, Paul O, et al. Alcohol, cardiovascular risk factors and mortality: the Chicago experience. *Circulation* 1981;64(suppl 3):20-7.
 63. Ross RK, Paganini-Hill A, Mack TM, et al. Menopausal oestrogen therapy and protection from death from ischaemic heart disease. *Lancet* 1981;1:858-60.
 64. Cullen K, Stenhouse NS, Wearne KL. Alcohol and mortality in the Busselton Study. *Int J Epidemiol* 1982;11:67-70.
 65. Kono S, Ikeda M, Ogata M, et al. The relationship between alcohol and mortality among Japanese physicians. *Int J Epidemiol* 1983;12:437-41.
 66. Metzner HL, Carman WJ, House J. Health practices, risk factors, and chronic disease in Tecumseh. *Prev Med* 1983;12:491-507.
 67. Kozarevic D, Vojvodic N, Gordon T, et al. Drinking habits and death: The Yugoslavia Cardiovascular Disease Study. *Int J Epidemiol* 1983;12:145-50.
 68. McGee DL, Reed DM, Yano K, et al. Ten-year incidence of coronary heart disease in the Honolulu Heart Program: relationship to nutrition intake. *Am J Epidemiol* 1984;119:667-76.
 69. Kromhout D, de Lezenne Coulander C. Diet, prevalence and 10-year mortality from coronary heart disease in 871 middle-aged men: The Zutphen Study. *Am J Epidemiol* 1984;119:733-41.
 70. Gordon T, Doyle JT. Drinking and coronary heart disease: The Albany Study. *Am Heart J* 1985;110:331-4.
 71. Friedman LA, Kimball AW. Coronary heart disease mortality and alcohol consumption in Framingham. *Am J Epidemiol* 1986;124:481-9.
 72. Ramsay LE. Alcohol and myocardial infarction in hypertensive men. *Am Heart J* 1979;98:402-4.
 73. Khaw KT, Barrett-Connor E. Dietary fiber and reduced ischemic heart disease mortality rates in men and women: a 12-year prospective study. *Am J Epidemiol* 1987;126:1093-102.
 74. Shaper AG, Wannamethee G, Walker M. Alcohol and mortality in British men: explaining the U-shaped curve. *Lancet* 1988;2:1267-73.
 75. Andreasson S, Allebeck P, Romelsjo A. Alco-

- hol and mortality among young men: longitudinal study of Swedish conscripts. *Br Med J* 1988;296:1021-5.
76. Rotevatn S, Akslen LA, Bjelke E. Lifestyle and mortality among Norwegian men. *Prev Med* 1989;18:433-43.
 77. Klatsky AL, Armstrong MA, Friedman GD. Risk of cardiovascular mortality in alcohol drinkers, ex-drinkers and nondrinkers. *Am J Cardiol* 1990;66:1237-42.
 78. Simons LA, Friedlander Y, McCallum J, et al. The Dubbo study of the health of elderly: correlates of coronary heart disease at study entry. *J Am Geriatr Soc* 1991;39:584-90.
 79. Suh I, Shaten J, Cutler JA, et al. Alcohol use and mortality from coronary heart disease: the role of high-density lipoprotein cholesterol. *Ann Intern Med* 1992;116:881-7.
 80. Klatsky AL, Armstrong MA, Friedman GD. Alcohol and mortality. *Ann Intern Med* 1992;117:646-54.
 81. Farchi G, Fidanza F, Mariotti S, et al. Alcohol and mortality in the Italian rural cohorts of the Seven Countries Study. *Int J Epidemiol* 1992;21:74-81.
 82. Division of Health Interview Statistics, National Center for Health Statistics. National Health Interview Survey 1983, alcohol supplement. (Data tape). Hyattsville, MD: National Center for Health Statistics, 1985.
 83. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am J Epidemiol* 1992;135:1301-9.
 84. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986;7:177-88.
 85. Berlin JA, Longnecker MP, Greenland S. Meta-analysis of epidemiologic dose-response data. *Epidemiology* 1993;4:218-28.
 86. Maclure M, Greenland S. Tests for trend and dose response: misinterpretations and alternatives. *Am J Epidemiol* 1992;135:96-104.
 87. Hastie T, Tibshirani R. Generalized additive models. New York: Chapman and Hall Ltd, 1990.
 88. Greenland S. A meta-analysis of coffee, myocardial infarction, and coronary death. *Epidemiology* 1993;4:366-74.
 89. Greenland S. Modelling and variable selection in epidemiologic analysis. *Am J Public Health* 1989;79:340-9.
 90. Longnecker MP, Berlin JA, Orza MJ, et al. A meta-analysis of alcohol consumption in relation to risk of breast cancer. *JAMA* 1988;260:652-6.
 91. Wilhelmsen L, Tibblin G, Elmfeldt D, et al. Coffee consumption and coronary heart disease in middle-aged Swedish men. *Acta Med Scand* 1977;201:547-52.
 92. Heyden S, Tyroler HA, Heiss G, et al. Coffee consumption and total mortality. *Arch Intern Med* 1978;138:1472-5.
 93. Murray SS, Bjelke E, Gibson RW, et al. Coffee consumption and mortality from ischemic heart disease and other causes: results from the Lutheran Brotherhood Study, 1966-1978. *Am J Epidemiol* 1981;113:661-7.
 94. Jacobsen BK, Bjelke E, Kvale G, et al. Coffee drinking, mortality, and cancer incidence: results from a Norwegian prospective study. *J Natl Cancer Inst* 1986;76:823-31.
 95. La Croix AZ, Mead LA, Liang KY, et al. Coffee consumption and the incidence of coronary heart disease. *N Engl J Med* 1986;315:977-82.
 96. Tverdal A, Stensvold I, Solvoll K, et al. Coffee consumption and death from coronary heart disease in middle aged Norwegian men and women. *BMJ* 1990;300:566-9.
 97. Rosengren A, Wilhelmsen W. Coffee, coronary heart disease and mortality in middle-aged Swedish men: findings from the Primary Prevention Study. *J Intern Med* 1991;230:67-71.
 98. Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results: follow-up of applications submitted to two institutional review boards. *JAMA* 1992;267:374-8.
 99. Stewart LA, Parmar MKB. Meta-analysis of the literature or of individual patient data: Is there a difference? *Lancet* 1993;341:418-22.
 100. Berlin JA, Colditz GA. A meta-analysis of physical activity in the prevention of coronary heart disease. *Am J Epidemiol* 1990;132:612-28.
 101. Hellerstedt WL, Jeffery RW, Murray DM. The association between alcohol intake and adiposity in the general population. *Am J Epidemiol* 1990;132:594-611.
 102. Matthews KA. Coronary heart disease and Type A behaviors: update on and alternative to the Booth-Kewley and Friedman (1987) quantitative review. *Psychol Bull* 1988;104:373-80.
 103. Folsom AR, Hughes JR, Buehler JF, et al. Do Type A men drink more frequently than Type B men? Findings in the Multiple Risk Factor Intervention Trial (MRFIT). *J Behav Med* 1985;8:227-35.
 104. Maclure KM, Hayes KC, Colditz GA, et al. Weight, diet, and the risk of symptomatic gallstones in middle-aged women. *N Engl J Med* 1989;321:563-9.
 105. Colsher PL, Wallace RB. Is modest alcohol consumption better than none at all? An epidemiologic assessment. *Annu Rev Public Health* 1989;10:203-19.
 106. Greenland S. Randomization, statistics and causal inference. *Epidemiology* 1990;1:421-9.
 107. Stehbens WE. Imprecision of the clinical diagnosis of coronary heart disease in epidemiological studies and atherogenesis. *J Clin Epidemiol* 1991;44:999-1006.
 108. Kannel WB, Abbott RD. Incidence and prognosis of unrecognized myocardial infarction: an update on the Framingham Study. *N Engl J Med* 1984;311:1144-7.
 109. Poole C. Exceptions to the rule about nondifferential misclassification. (Abstract). *Am J Epidemiol* 1985;122:508.
 110. Goldman L, Cook EF, Brand DA, et al. A computer protocol to predict myocardial infarction in emergency department patients with chest pain. *N Engl J Med* 1988;318:797-803.
 111. Narula J, Khaw BA, Dec GW Jr, et al. Brief

- report: recognition of acute myocarditis masquerading as acute myocardial infarction. *N Engl J Med* 1993;328:100-4.
112. Veenstra J. Effects of alcohol on the cardiovascular system. In: Veenstra J, van der Heij DG, eds. Alcohol and cardiovascular disease: proceedings of an international symposium, Scheveningen, Netherlands, 1-2 October 1991. Wageningen, The Netherlands: Centre for Agriculture Publishing and Documentation (PUDOC), 1992:87-113.
 113. Cobb LA, Baum RS, Alvarez H III, et al. Resuscitation from out-of-hospital ventricular fibrillation: 4 years follow-up. *Circulation* 1975;51(suppl 3):223-35.
 114. Poikolainen K. Underestimation of recalled alcohol intake in relation to actual consumption. *Br J Addict* 1985;80:215-16.
 115. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol* 1985;122:51-65.
 116. Rimm EB, Giovannucci EL, Stampfer MJ, et al. Reproducibility and validity of an expanded self-administered semiquantitative food frequency questionnaire among male health professionals. *Am J Epidemiol* 1992;135:1114-26.
 117. Rosner B, Spiegelman D, Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error. *Am J Epidemiol* 1990;132:734-45.
 118. Apparent per capita ethanol consumption—United States, 1977-1986. *MMWR* 1989;38:800-3.
 119. Shaper AG. Alcohol and mortality: a review of prospective studies. *Br J Addict* 1990;85:837-47.
 120. Gordon T, Kannel WB, Dawber TR, et al. Changes associated with quitting cigarette smoking: The Framingham study. *Am Heart J* 1975;90:322-8.
 121. Le Marchand L, Kolonel LN, Hankin JH, et al. Relationship of alcohol consumption to diet: a population-based study in Hawaii. *Am J Clin Nutr* 1989;49:567-72.
 122. Colditz GA, Giovannucci E, Rimm EB, et al. Alcohol intake in relation to diet and obesity in women and men. *Am J Clin Nutr* 1991;54:49-55.
 123. Ulbricht TLV, Southgate DAT. Coronary heart disease: seven dietary factors. *Lancet* 1991;338:985-92.
 124. Rimm EB, Stampfer MJ, Ascherio A, et al. Vitamin E supplementation and the risk of coronary heart disease in men. *N Engl J Med* 1993;328:1450-6.
 125. Stampfer MJ, Hennekens CH, Manson JE, et al. Vitamin E consumption and the risk of coronary disease in women. *N Engl J Med* 1993;328:1444-9.
 126. Puccio EM, McPhillips JB, Barrett-Connor E, et al. Clustering of atherogenic behaviors in coffee drinkers. *Am J Public Health* 1990;80:1310-13.
 127. Hennekens CH, Peto R, Hutchison GB, et al. An overview of the British and American aspirin studies. *N Engl J Med* 1988;318:923-4.
 128. Stampfer MJ, Colditz GA, Willett WC, et al. Postmenopausal estrogen therapy and cardiovascular disease: ten-year follow-up from the Nurses' Health Study. *N Engl J Med* 1991;325:756-62.
 129. Robins J. The control of confounding by intermediate variables. *Stat Med* 1989;8:679-701.
 130. Stampfer MJ, Colditz GA, Willett WC, et al. A prospective study of moderate alcohol drinking and risk of diabetes in women. *Am J Epidemiol* 1988;128:549-58.
 131. Hill AB. The environment and disease: association or causation? *Proc R Soc Med* 1965;58:295-300.
 132. Kamarck T, Jennings JR. Biobehavioral factors in sudden cardiac death. *Psychol Bull* 1991;109:42-75.
 133. Benfante RJ, Reed DM, MacLean CJ, et al. Risk factors in middle age that predict early and late onset of coronary heart disease. *J Clin Epidemiol* 1989;42:95-104.
 134. Veenstra J. Effects of alcohol on blood lipids and platelet function. In: Veenstra J, van der Heij DG, eds. Alcohol and cardiovascular disease: proceedings of an international symposium, Scheveningen, Netherlands, 1-2 October 1991. Wageningen, The Netherlands: Centre for Agricultural Publishing and Documentation (PUDOC), 1992:65-74.
 135. van der Heij DG, Schaafsma G, eds. Biomedical and social aspects of alcohol use: a review of the literature. Wageningen, The Netherlands: Centre for Agricultural Publishing and Documentation (PUDOC), 1991.
 136. Stampfer MJ, Sacks FM, Salvini S, et al. A prospective study of cholesterol, apolipoproteins, and the risk of myocardial infarction. *N Engl J Med* 1991;325:373-81.
 137. Maclure M, Hankinson S. Analysis of selection bias in a case-control study of renal adenocarcinoma. *Epidemiology* 1990;1:441-7.
 138. Stampfer MJ, Rimm EB, Walsh DC. Commentary: alcohol, the heart and public policy. *Am J Public Health* 1993;83:801-4.
 139. Manson JE, Tosteson H, Ridker PM, et al. The primary prevention of myocardial infarction. *N Engl J Med* 1992;326:1406-16.
 140. Thompson SG, Pocock SJ. Can meta-analyses be trusted? *Lancet* 1991;338:1127-30.
 141. Greenland S. Critical look at some popular meta-analytic methods. (Abstract). *Am J Epidemiol* 1993;138:672.
 142. Petitti D. Meta-analysis of non-experimental studies: problems and prospects. (Abstract). *Am J Epidemiol* 1993;138:672-3.
 143. Shapiro S. Meta-analysis/shmeta-analysis. (Abstract). *Am J Epidemiol* 1993;138:673.