Case Study Summary

Title: Applying Hypothesis-Testing Methods to Help Inform Causality Conclusions from Epidemiology Studies

Version: 1

Presented by: Sabine Lange, Lalita Shrestha, Swati Rawat

Panel Advisor:

1. **Provide a few sentences summarizing the method illustrated by the case study.**

The application of epidemiological or observational data to dose-response evaluations in risk assessment requires that there is sufficient proof that the chemical of interest *caused* the relevant adverse effect. However, because epidemiological studies are associational, they cannot prove causation. The following case study explores the question of *whether we can use a hypothesis-testing paradigm for evaluating data from epidemiology studies to better judge whether the study results are likely to represent a causal relationship*. The case describes investigating patterns in the epidemiology results based on: 1) exposure and outcome variability; 2) specificity of the health effect; 3) exposure concentration (i.e. dose-response); and 4) severity of the health effect. First, we interrogated the literature and conducted simple simulation studies to ensure that the patterns were conceptually sound and could appropriately be applied to epidemiology study results. This step revealed that exposure and/or outcome variability may not provide a consistent pattern of epidemiology study results, depending on the complexity of the analysis. The next question was whether these patterns can actually be observed in epidemiology studies and whether they appropriately correlate with the final causal conclusion. If these patterns are not consistent with causality, then the pressing question becomes: why not? We tested for the presence of patterns of dose-response and outcome specificity using a positive control scenario (cigarette smoking and lung cancer), and a negative control scenario (vitamin supplementation and cancer). We found that 6 positive control studies confirm that two of our hypothesis-testing patterns (dose-response and outcome specificity) can be found in a known causal relationship. The 9 negative control studies demonstrated a lack of consistent pattern for dose-response or outcome specificity for three of the four investigated nutrient-outcome pairs. These results show promise for using these patterns to identify appropriate causal conclusions, but further work is required to determine the presence of dose-response and outcome specificity patterns for one of the negative control nutrient-outcome pairs.

2. **Describe the problem formulation(s) the case study is designed to address. How is the method described in the case useful for addressing the problem formulation?**

Epidemiology or observational studies provide a crucial piece of information in a risk assessment – that is, information about how a chemical affects human beings. In some cases, these studies allow investigation of populations and chemical concentrations that otherwise would have no available information – such as vulnerable populations, and with lower exposure concentrations. They are also conducted outside of controlled environments, providing information that could be more relevant to actual human exposure. Unfortunately, the application of these studies for risk assessment purposes – either for hazard assessment or dose-response – requires an answer to the question: did the chemical *cause* the health effect? This case study investigates a method that may help address that question.

This case study addresses the question of causality through the application of a hypothesis-testing paradigm that is applied to the epidemiology literature available for a specific chemical and health endpoint. The concept is that, if a true causal association exists between a chemical exposure and a health effect, then certain patterns of results may be found, based on:

1) Exposure and outcome variability;
2) Specificity of the health effect;
3) Exposure concentration (i.e. dose-response); and
4) Severity of the health effect.

In this work, we conducted a conceptual evaluation of each of these patterns based on the literature and simulation studies, and then we tested them on positive and negative control epidemiology study results. The purpose of using positive and negative controls is to see if these patterns could be used to predict the ultimate causal conclusions from these associational studies – i.e. were the patterns present for relationships that turned out to be causal, and not present for relationships that turned out not to be causal.

Concept and Simulation Studies

Exposure and Outcome Variability

The hypothesis: It has often been stated that when there is an increase in the random variability or mis-classification of the exposure, then the effect estimate will be attenuated (aka biased towards the null) (Hausman, 2001). There are two major types of exposure error: classical and Berkson (Zeger et al., 2000). With classical error the exposure estimate varies randomly around the true value and has a greater variation than the true value (e.g. instrument measurement error). For Berkson error the true value varies randomly around the estimate and has greater variation than the estimated values (e.g. an exposure estimate that is based on an average of many instrument measurements). In a simple system classical error is expected to attenuate the risk estimate, whereas Berkson error should not bias the effect estimate but will increase its variability. Figure 1 shows a simple simulation study of data with and without random variability in the x (exposure) variable, demonstrating a decrease in the slope of the line (from 1.05 to 0.83) with increased variability.

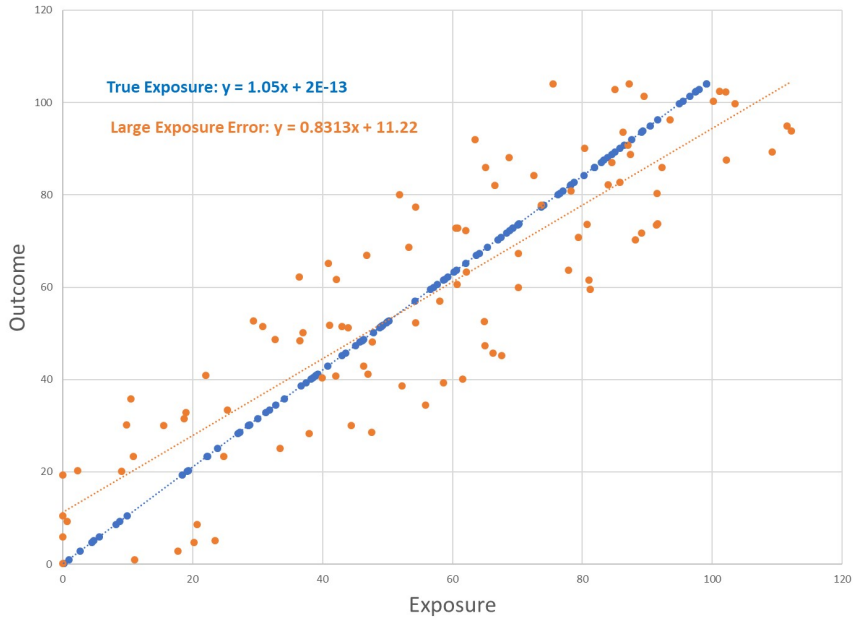

Figure 1. Decreased slope of relationship between exposure and outcome with increased variability in the exposure variable.

Therefore, in general, one would hypothesize that in two similar studies, the study with the more precise exposure measurement should have a higher effect estimate. Further investigation into the literature demonstrates that the reality is a bit more complicated. For example, in a simulation study Szpiro et al., (2011) found that the above hypothesis is generally true (better exposure specification led to narrower confidence intervals and higher effect estimates). However, they also found that the health effect estimate was not changed, but the confidence intervals were narrower, in models with mis-specified vs correctly specified exposures where a missing covariate has less variability than other covariates. Further, studies have shown that classical error can bias effect estimates towards the null and Berkson error can widen confidence intervals in a simple, single pollutant model where (1) the concentration-response is genuinely linear (Fuller, 1987), (2) measured concentrations are good surrogates for personal exposure, and (3) differences between the measured and the personal exposures are constant (Zeger et al., 2000). Further, multiple pollutants are often modeled to consider confounding effects, but classical error using multiple linear regression models can bias towards or away from the null (Zeger et al., 2000) because of the interplay between interpollutant correlations and the measurement error for each pollutant (Carrothers and Evans, 2000). Many studies have been conducted that show that this relationship is actually quite complicated, and unless the study has a very simple, one-variable linear analysis (Brakenhoff et al., 2018; Hausman, 2001; Jurek et al., 2008, 2005; Loken and Gelman, 2017), one should not make an assumption of effect estimate attenuation with increasing exposure error.

In contrast to exposure measurement error, in a simple system outcome mismeasurement will not attenuate the effect estimate (see simulation in Figure 2) but will increase the uncertainty around the estimate (Hausman, 2001). However, if there is a limitation on the outcome, such

as the use of a logit or probit estimation (where the outcome can only be zero or one), then the effect estimate can be biased or show different patterns with increasing error.

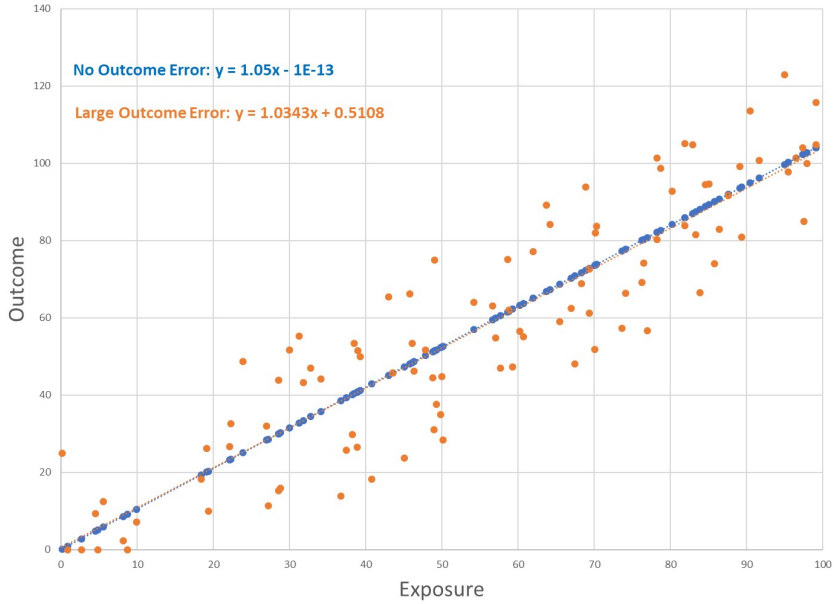

Figure 2. No change in slope of relationship between exposure and outcome with increased variability in the outcome variable in a simple linear regression.

These studies demonstrate that the effect of outcome or exposure error on an effect estimate is not necessarily simple to predict. Therefore, we did not use outcome or exposure error as a predictor in our hypothesis test of causality of an epidemiological association.

Outcome Specificity

The hypothesis: Specific health endpoints that are causally related to the exposure should have greater effect estimates than more general endpoints that include both causally-related and non-causally-related endpoints. For example, if an agent increases liver cancer specifically, then the effect estimate for liver cancer should be higher than the effect estimate for all cancers combined. The principle here is a straight forward signal-to-noise problem. If there is only one causal relationship in a subset of a dataset, but many other data points are included that are not causally related to the exposure, then there will be a diminishment of the signal from the subset of data. This is demonstrated by a simulation study in Figure 3 – showing both a clean dataset (A), and a dataset with additional error in the variables (B).

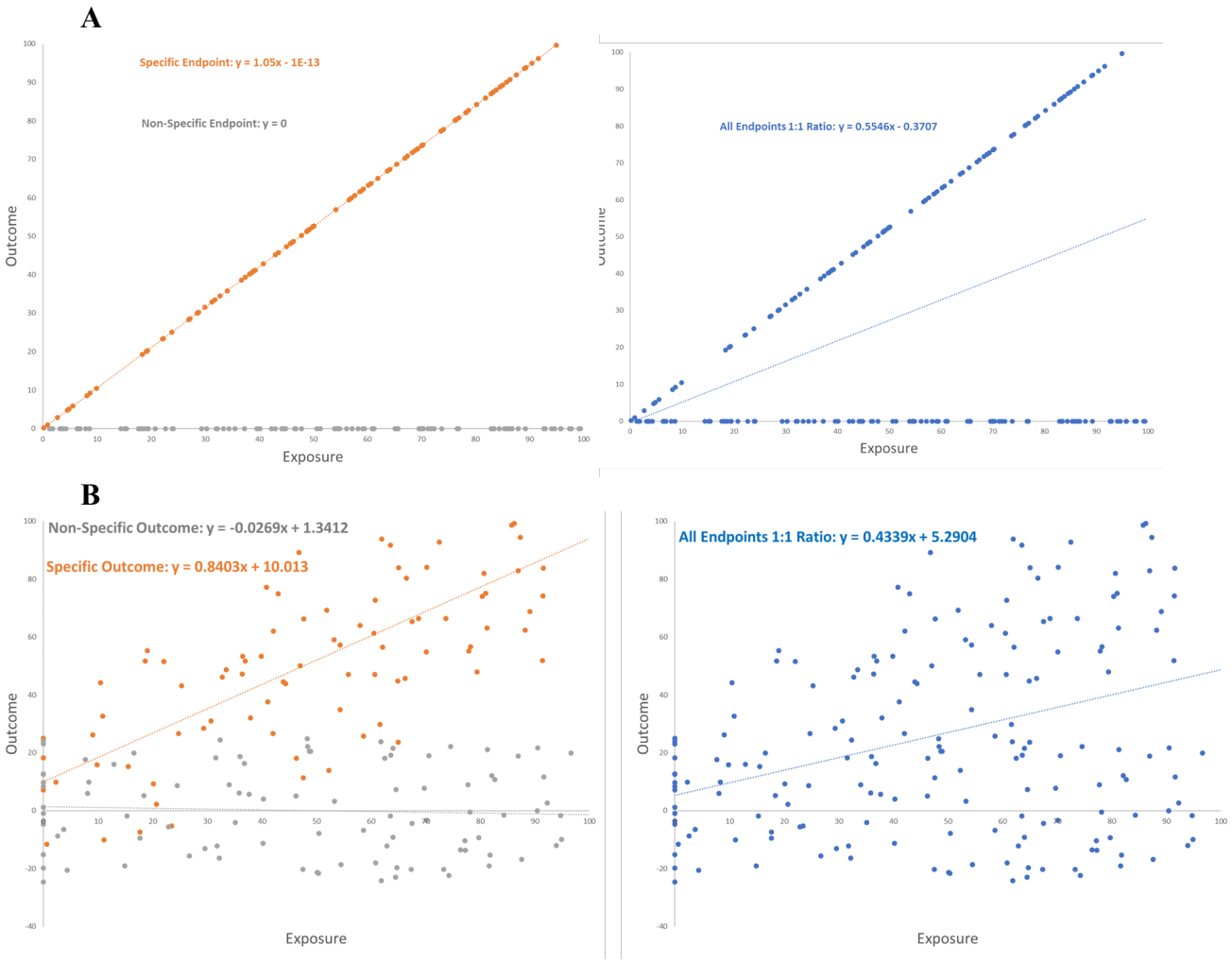

Figure 3. Decrease in slope of relationship between exposure and outcome when an unrelated outcome is included in the analysis of either clean data (A) or data with error in the exposure and outcome variables (B).

Based on the simulation studies, several patterns can be seen in the data:

- Added exposure error or outcome error doesn't change the difference in the effect estimate generated using total versus specific outcomes
- Having a small effect estimate does not seem to decrease the difference in the effect estimate generated using total versus specific outcomes
- The higher the number of non-specific outcome endpoints, the greater the difference between specific and non-specific effect estimates
- All these patterns are true for both linear and log-linear regression analyses

For our hypothesis-testing analysis, we would look for studies that investigate both more and less specific endpoints in the same group (e.g. all cancer and lung cancer, all mortality and CVD mortality, etc). Based on the specificity hypothesis, we would expect that if a specific endpoint is genuinely causally related to the exposure, that it should have a higher effect

estimate compared to the larger, less specific endpoint group (unless we expect all the effects in the less specific endpoint group to be causally related to the exposure). In addition, the less common the specific outcome, the greater the expected differential compared to the effect estimate for the larger endpoint group.

Dose (Exposure) - Response

The hypothesis: Based on toxicological theory, higher exposure concentrations should produce greater effect estimates. Often epidemiology studies present a single effect estimate (a slope, relative risk, odds ratio, hazard ratio, etc) to represent the relationship between exposure and outcome. From a dose-response (or exposure-response) standpoint, there are several ways to interpret this:

- If the effect estimate is statistically significant, this estimate demonstrates the presence of a dose-response between exposure and outcome
- In the absence of the primary data, dose-response cannot be assessed because the model assumes a certain shape and a constant increase in outcome with dose (the model or study does not necessarily test whether this assumption is valid for the relationship between the exposure and the outcome)

One way to test for the presence of a dose (exposure)-response is to look at categorical results – a dose (exposure)-response would demonstrate an increasing effect estimate with increasing dose, relative to a single reference group. As shown in our simulation studies, this relationship is true for both linear and log-linear relationships, and with exposure and/or outcome error in the data.

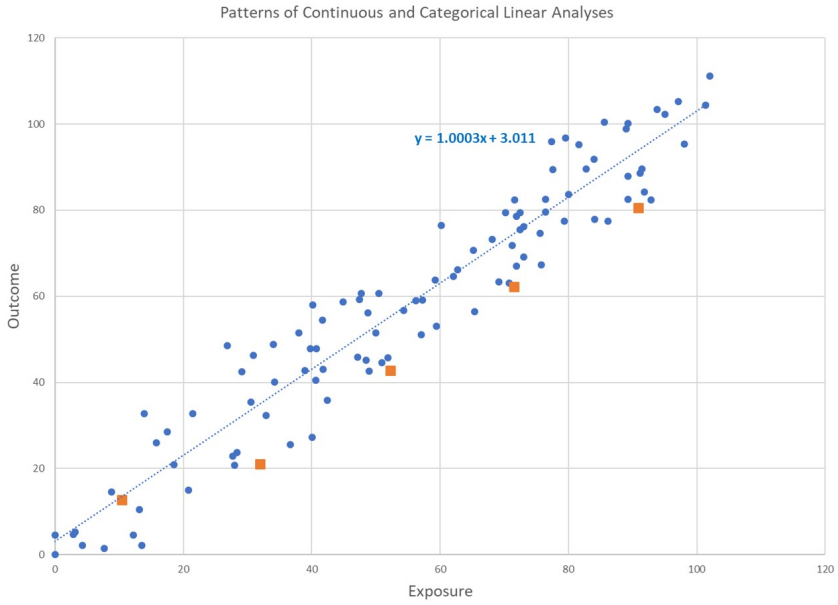

Figure 4. Clear exposure-response relationship in quintiles analysis based on data with an underlying exposure response. Blue dots are the continuous data (with the equation for the relationship), orange squares are the categorical data points showing increasing effect compared to the first quintile.

For our hypothesis-testing analysis, we would look for studies that investigate categorical dose or exposure-response for a single endpoint. Based on the dose-response hypothesis, we would expect increasing effect estimates for increasing dose, when compared to a single reference group.

Outcome Severity

The hypothesis: Based on toxicological theory, we expect that higher exposure concentrations should produce more severe health effects than lower exposure concentrations; and that at the same exposure concentrations we should observe larger and less variable effect estimates for less severe health endpoints compared to more severe health endpoints. Based on this theory, for linear and log-linear relationships without modeled thresholds, the slope of the relationship (and therefore the hazard ratio or relative risk) would be expected to be higher for less severe effects compared to more severe effects. This theory is based on the concept of the risk of an effect in the population. For example, let's say that at 100 units of exposure, a low severity effect has a risk of 0.5, a moderately severe effect has a risk of 0.3, and a highly severe effect has a risk of 0.1, and that for 0 units of exposure there is 0 risk. The slopes of these lines would be 0.005, 0.003, and 0.001, respectively.
We confirmed this concept with simple simulation studies as shown in Figure 5. The relationship holds true with random error incorporated into the exposure or outcome variables. Because this concept is based on the probability of an effect in the population, the severity hypothesis may not apply to case-control studies (that generate odds and odds ratios) because these studies do not provide probability information about the total population, just for the case and control groups.

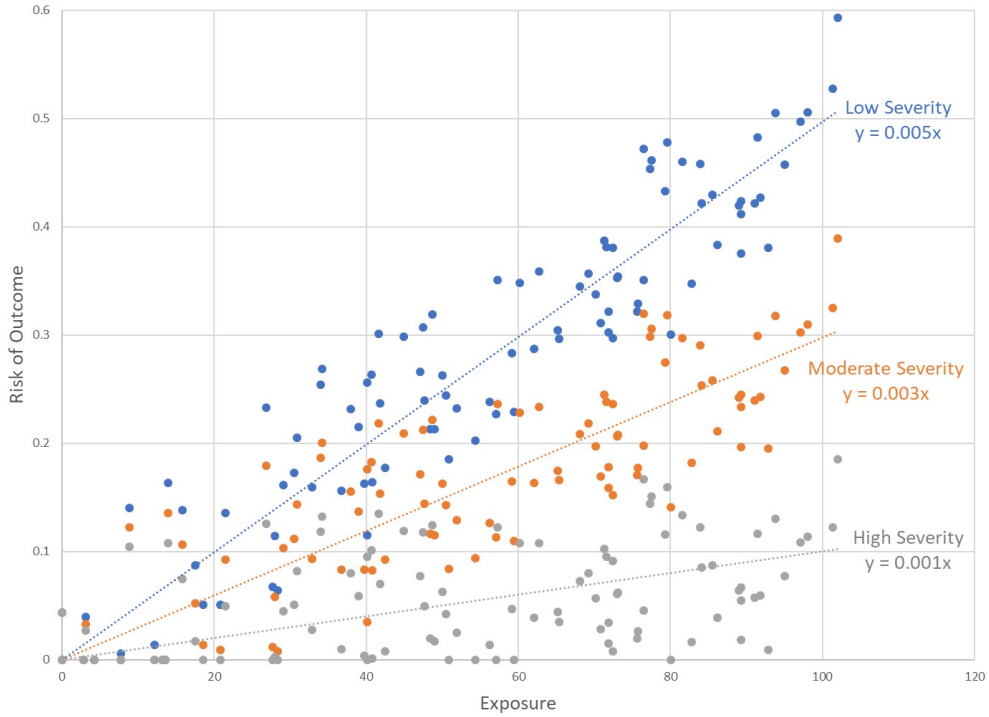

Figure 5. Higher slopes of effect with decreasing outcome severity in the same exposure concentration range. Underlying probabilities of effect in the simulation were per 100 units: 0.5 for low severity, 0.3 for moderate severity, and 0.1 for high severity.

For our hypothesis-testing analysis, we would look for studies that investigate different severities of health endpoints (that occur in the same health effect pathway, such as asthma exacerbations and asthma hospital admissions). Based on the severity hypothesis, we would expect increasing effect estimates for less severe endpoints, compared to more severe endpoints.

Positive Control: Smoking and Lung Cancer

The positive control for this case study is cigarette smoking and lung cancer. The causal relationship between smoking and lung cancer has been definitively established, which makes it an excellent positive control to test for the presence of our hypothesized patterns in epidemiological data.

We looked at patterns of dose-response and/or outcome specificity in 6 studies. These studies included cohort and case-control studies that investigated the associations between smoking and cancer.

For dose-response, the data from three studies (Freedman et al., 2008; Powell et al., 2013; Remen et al., 2018), using two different risk metrics (odds ratio, hazard ratio), as well as with multiple types of dose metrics, demonstrated clear dose-response of smoking with lung cancer (Table 1). Using the data from The Health Improvement Network (THIN), a UK medical research database, a data set comprising 12,121 incident cases of lung cancer and 48,216 age-, sex-, and general practice-matched control subjects, Powell et al. (2013) studied the association between smoking quantity and lung cancer in men and women. Conditional logistic regression was used to calculate odds ratios for lung cancer according to highest-ever quantity smoked in men and women separately. An increase in smoking quantity was associated with an increase of risk estimates regardless of the sex of the subjects. Similarly, Remen et al. (2018) conducted a case-control study to provide new data on risk of lung cancer in relation to various metrics of smoking history. This study included 1,203 lung cancer cases and 1,513 age-, gender-, and residential area-matched controls residing in Montreal and its surrounding suburbs. Regardless of the smoking metrics used, increase in smoking quantity was associated with an increase in the lung cancer risk estimate. Similar findings were reported by Freedman et al. (2008), where the authors investigated susceptibility of men and women to cigarette smoking by comparing lung carcinoma incidence rates by stratum from the National Institutes of Health – AARP cohort. The study included 279,214 men and 184,263 women from eight US states aged 50 to 71 years. Smoking was associated with increased lung carcinoma risk in both men and women.

Three additional studies (Lewer et al., 2017; Ordóñez-Mena et al., 2016; Thun et al., 2013) provided information about outcome specificity, and clearly demonstrated a greater slope of response for the association between smoking and lung cancer incidence or mortality (or chronic obstructive pulmonary disease, or head and neck cancer), compared to the

association between smoking and total cancer incidence or total mortality (Table 2). Lewer et al. (2017) tested the hypothesis that smoking confers a greater mortality risk for individuals in low socioeconomic groups using a cohort of 18,479 adults drawn from the English Longitudinal Study of Aging (ELSA). Out of 5,050 deaths, 310 were due to lung cancer and 274 deaths were due to COPD. Thun et al. (2013) studied the temporal trends in mortality associated with cigarette smoking across three time periods; 1959-1965, 1982-1988, and 2000-2010, comparing absolute and relative risks according to sex and self-reported smoking status in two historical cohort studies and in five pooled contemporary cohort studies. The study concluded that the risk of death from cigarette smoking continues to increase among women and the increased risks are nearly identical for men and women in the contemporary cohort as compared with persons who have never smoked. Ordóñez-Mena et al. (2016) comprised a meta-analysis of 19 population-based prospective cohort studies with individual data for 897,021 European and American adults. The authors estimated rate advancement periods (RAPs) for the association of smoking exposure with total and site-specific (lung, breast, colorectal, prostate, gastric, head and neck, and pancreatic cancer) cancer incidence and mortality. This investigation showed that smoking considerably advances the risk of developing and dying from cancer.

In addition, one of the studies (Ordóñez-Mena et al., 2016) provides information that can be used to test the severity hypothesis: i.e. that more severe endpoints would have lesser associations than less severe endpoints, at the same exposure concentrations. If we consider cancer incidence to be a less severe, but related, endpoint to cancer mortality, then we would expect a greater association between smoking and incidence, than between smoking and mortality. This pattern is seen with lung cancer, but not with total cancers or head-and-neck cancers. This hypothesis-testing endpoint will need to be investigated further, with other studies.

The 6 available studies confirm that two of our hypothesis-testing patterns can be found in a known causal relationship.

Table 1. Summary of Dose-Response Analysis of Studies Investigating the Association between Smoking and Lung Cancer

Exposure Variable	Ref Quantile	Quantile 2	Quantile 3	Quantile 4	Quantile 5
Powell et al., 2013, Lung Cancer, Odds Ratios (95% CI)					
Smoking Quantity	**Never**	**Light**	**Moderate**	**Heavy**	
	1	9.32 (8.48-10.25)	11.78 (10.79-12.87)	15.02 (13.69-16.48)	
Remen et al., 2018, Lung Cancer in Women, Odds Ratios (95% CI)					
Duration of Smoking (yrs)	**Never (0)**	**0-20**	**20-30**	**30-40**	**> 40**
	1	1.51 (0.74-3.04)	6.37 (3.55-11.41)	13.64 (8.19-22.74)	28.79 (16.86-49.16)
Intensity of	**0**	**0-20**	**20-30**	**> 30**	

Smoking (cig/day)	1	6.05(3.70-9.90)	19.4(11.81-31.86)	18.2(10.10-32.80)	
Pack-years	**0**	**0-20**	**20-40**	**40-60**	**> 60**
	1	2.04 (1.11-3.74)	8.66 (5.10-14.68)	25.48 (15.08-43.04)	37.39 (19.79-70.62)
Cumulative Smoking Index (CSI)[a]	**0**	**0 – 1**	**1 – 2**	**> 2**	
	1	1.25 (0.62-2.51)	11.98 (7.32-19.62)	29.66 (17.67-49.80)	
Freedman et al., 2008, Lung Cancer in Current Smokers, Hazard Ratios (95% CI)					
Intensity of Smoking (cig/day) in Men	**Never (0)**	**1-10**	**11-20**	**21-30**	**31-40**
	1	20.7 (16.3-26.3)	30.5 (24.6-37.9)	35.9 (28.7-44.8)	42.6 (33.8-53.8)
Intensity of Smoking (cig/day) in Women	1	13.4 (10.9-16.5)	22.5 (18.8-27.1)	25.2 (20.5-31.0)	40.7 (32.3-51.2)

[a]Cumulative Smoking Index (CSI) takes into consideration duration of smoking, time since cessation, and daily amount of cigarettes smoked.

Table 2. Summary of Outcome Specificity Analysis of Studies Investigating the Association between Smoking and Mortality

Exposure Variable	**Less Specific Outcome**	**More Specific Outcomes**				
Lewer et al., 2017, Mortality, *Ratio* of age-adjusted mortality rate per 100,000 person years in smokers : never-smokers						
	Never Smoker	**Ex-Smoker**	**Current Smoker**			
All-Cause Mortality	1	1.33	1.87			
No. Cases	2059	2748	3842			
Lung Cancer or COPD	1	5.17	13.25			
No. Cases	60	310	795			
Thun et al., 2013, Mortality, Relative Risk of mortality among those 55-years or older, for current smokers compared to never smokers, 2000-2010 (95% CI)						
	All Cause Mortality	**Lung Cancer**	**COPD**			
Women	2.76 (2.69 - 2.84)	25.66 (23.17-28.40)	22.35 (19.55-25.55)			

No. Cases	62965	4785	3034			
Men	2.80 (2.72-2.88)	24.97 (22.20-28.09)	25.61 (21.68-30.25)			
No. Cases	73800	6635	3478			
Ordóñez-Mena et al., 2016, Cancer Incidence or Mortality, Hazard ratio of current smokers compared to never-smokers (95% CI)						
	Total Cancer	**Lung Cancer**	**Head and Neck Cancer**	**Colorectal Cancer**	**Breast Cancer**	**Prostate Cancer**
Cancer Incidence	1.44 (1.28-1.63)	13.1 (9.90-17.3)	2.89 (1.98-4.21)	1.20 (1.07-1.34)	1.07 (1.00-1.15)	0.81 (0.72-0.91)
No. Cases	26007	6333	1051	2064	2536	3701
Cancer Mortality	2.19 (1.83-2.63)	11.5 (8.21-16.1)	3.74 (2.38-5.89)	1.35 (1.16-1.58)	1.28 (1.06-1.55)	1.26 (0.97-1.64)
No. Cases	13450	6165	359	912	466	589

Negative Control: Dietary Constituents and Cancer

The negative control for this case study is the association observed between various dietary constituents in the serum and cancer incidence or mortality. Multiple epidemiology studies conducted in the 1980s and 1990s demonstrated associations between, among others, β-carotene serum concentrations and lung cancer, retinol (vitamin A) and lung cancer, and α-tocopherol (vitamin E) and lung and prostate cancer.

Subsequently the relationship between β-carotene and retinol and lung cancer was tested in the Beta-Carotene and Retinol Efficacy Trial (CARET), a randomized, controlled clinical trial of 18,314 men and women (Omenn et al., 1996). The subjects in this trial received either a placebo or a combination of 30 mg β-carotene plus 25000 IU retinyl palmitate (vitamin A). The trial was terminated 21 months earlier than expected because it showed increased risk of lung cancer incidence and mortality among the intervention group. Similarly, the Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) Study enrolled 29,133 male smokers into a randomized, double-blind, placebo-controlled primary prevention trial to determine whether supplementation with α-tocopherol (vitamin E), β-carotene, or both, would reduce the incidence of lung cancers or other cancers (Alpha-Tocopherol, 1994). α-Tocopherol supplementation did not have any effect on lung cancers but showed a decrease in prostate cancer (later shown not to be a causal relationship in the SELECT trial, Klein et al., 2011). As with the CARET trial, β-carotene supplemented-men had an increased rate of lung cancer incidence.

We looked at patterns of dose-response and/or outcome specificity in 9 studies. These studies included cohort and case-control studies that investigated the associations between β-carotene, retinol, and/or α-tocopherol and cancer incidence or mortality. The studies largely used a matched case-control design with matching for sex, age, smoking status, and a variety of other factors.

Dose-Response Patterns:

- β-Carotene: serum concentration showed a positive dose-response with lung cancer in some, but not all, of the studies (Table 3).
- Retinol: the two studies of lung cancer and retinol did not show evidence of dose-response (Table 4).
- α-Tocopherol: there was marginal evidence for a dose-response with α-tocopherol and lung cancer, but not with prostate cancer (Table 5).

Outcome Specificity Patterns:

- β-Carotene: most of the studies show a greater slope of association between serum β-carotene and lung cancer, compared to the slope between serum β-carotene and total cancer (Table 6).
- Retinol: there is conflicting evidence of a greater slope of association between serum retinol and lung cancer, compared to the slope between serum retinol and total cancer (Table 7).
- α-Tocopherol: there is conflicting evidence of a greater slope of association between serum α-tocopherol and lung or prostate cancer, compared to the slope between serum α-tocopherol and total cancer (Table 8).

Data from these observational studies demonstrates that there is not a consistent pattern of dose-response or outcome specificity for the association between retinol and lung cancer, or α-tocopherol and either lung or prostate cancer. This lack of pattern was a likely clue that there was not a causal relationship between these vitamins and cancer. In contrast, there is some evidence for patterns of dose-response and outcome specificity between lower serum β-carotene concentrations and increased lung cancer. This is particularly interesting given results from multiple randomized controlled trials demonstrating the presence of the opposite relationship. This will require further study.

Table 3. Summary of Dose-Response Analysis of Studies Investigating the Association between Beta-Carotene and Cancer (lowest concentration hypothesized to be highest risk)

Exposure Variable	**Ref Quantile**	**Quantile 2**	**Quantile 3**	**Quantile 4**	**Quantile 5**	**Trend p-value**
Comstock et al., 1991, Lung Cancer Incidence, Odds Ratio						
Serum Concentration	**1st (Highest)**	**2nd**	**3rd**	**4th**	**5th (Lowest)**	**Trend**
	1	1.2	1.8	1.7	2.2	0.04
Connett et al., 1989, Lung Cancer Deaths, Odds Ratio						
Serum Concentration	**1st (Highest)**	**2nd**	**3rd**	**4th**	**5th (Lowest)**	**Trend**
B-Carotene	1	2.17	2.72	1.6	2.32	0.08
Nomura et al., 1985, Lung Cancer Incidence, Odds Ratio (95% CI)						
Serum Concentration	**1st (Highest)**	**2nd**	**3rd**	**4th**	**5th (Lowest)**	**Trend**
Unadjusted estimate	1	1.7 (0.6-4.7)	1.5 (0.5 - 4.1)	2.9 (1.1-7.3)	3.4 (1.4 - 8.4)	0.004

Adjusted estimate[a]	1	1.5 (0.5 - 4.1)	1.2 (0.4 - 3.5)	2.4 (0.9 - 6.2)	2.2 (0.8 - 6)	0.04

[a] adjusted for age and smoking using multiple logistic regression

Table 4. Summary of Dose-Response Analysis of Studies Investigating the Association between Vitamin A/Retinol and Cancer (lowest concentration hypothesized to be highest risk)

Exposure Variable	Ref Quantile	Quantile 2	Quantile 3	Quantile 4	Quantile 5	Trend p-value
Friedman et al., 1986, Lung Cancer Incidence, Odds Ratio						
Serum Concentration	**1st (Highest)**	**2nd**	**3rd**	**4th**	**5th (Lowest)**	**Trend**
Unmatched analysis	1	1.4	1.1	0.9	1.2	
Matched analysis	1	1.3	1.1	0.9	1.2	
Menkes et al., 1986, Lung Cancer Incidence, Odds Ratio						
Serum Concentration	**1st (Highest)**	**2nd**	**3rd**	**4th**	**5th (Lowest)**	**Trend**
	1	1.62	0.73	0.92	1.13	0.68

Table 5. Summary of Dose-Response Analysis of Studies Investigating the Association between Vitamin E/α-Tocopherol and Cancer (lowest concentration hypothesized to be highest risk)

Exposure Variable	Ref Quantile	Quantile 2	Quantile 3	Quantile 4	Quantile 5	Trend p-value
Comstock et al., 1991, Cancer Incidence, Odds Ratio						
Serum Concentration	**1st (Highest)**	**2nd**	**3rd**	**4th**	**5th (Lowest)**	**Trend**
Lung	1	1.3	2.2	1.9	2.5	0.04
Prostate	1	1.6	1.4	1.1		0.94
Menkes et al., 1986, Lung Cancer Incidence, Odds Ratio						
Serum Concentration	**1st (Highest)**	**2nd**	**3rd**	**4th**	**5th (Lowest)**	**Trend**
	1	1.32	2.19	1.87	2.48	0.04

Table 6. Summary of Outcome Specificity Analysis of Studies Investigating the Association between Beta-Carotene and Cancer

Exposure Variable	Less Specific Outcome	More Specific Outcomes				
Wald et al., 1988, Cancer Incidence, Percent Difference in Serum Concentration between Cases and Controls (Std Error)						
Serum	**All**	**Lung**	**Colorectal**	**Stomach**	**Bladder**	**CNS**

Concentration	**Cancer**					
	- 10% (4)	- 22% (8)	- 11% (10)	-27 % (17)	- 9% (20)	- 10% (15)
No. Cases	271	50	30	13	15	17
Connett et al., 1989, Cancer Deaths, Mean Difference in Serum Concentration between Cases and Controls						
Serum Concentration (µg/dL)	**All Cancer**	**Lung**	**Colon**	**GI Tract**	**Bladder & Kidney**	
B-Carotene	-0.6	-2.70	-1	0.6	2	
No. Cases	156	66	14	28	7	
Knekt et al., 1990, Cancer Incidence, Mean Difference in Serum Concentration between Cases and Controls (Std Error)						
Serum Concentration (µg/L)	**All Cancer**	**Lung**	**Colon**	**Rectum**	**Stomach**	**Prostate/ Breast**
Men	-11.8	-17.2	-10.8	16	-7.4	-0.4
No. Cases	453	144	6	15	48	37
Women	-7	-40	10.8	-37.3	27	-18.8
No. Cases	313	8	13	22	28	67
Willett et al., 1984, Cancer Incidence, *Carotenoids*, Mean Difference in Serum Concentration between Cases and Controls (Std Error)						
Serum Concentration (µg/L)	**All Cancer**	**Lung**	**Breast**	**Prostate**	**GI**	
	8.2 (6.4)	9 (16.5)	8.9 (17.2)	4.3 (19.4)	10.5 (19.9)	
No. Cases	111	17	14	11	11	

Table 7. Summary of Outcome Specificity Analysis of Studies Investigating the Association between Vitamin A/Retinol and Cancer

Exposure Variable	**Less Specific Outcome**	**More Specific Outcomes**				
Knekt et al., 1990, Cancer Incidence, Mean Difference in Serum Concentration between Cases and Controls (Std Error)						
Serum Concentration (µg/L)	**All Cancer**	**Lung**	**Colon**	**Rectum**	**Stomach**	**Prostate/ Breast**
Men	-22	-38	-93	-41	-19	16
No. Cases	453	144	6	15	48	37
Women	-17	24	-25	-30	4	-22
No. Cases	313	8	13	22	28	67
Willett et al., 1984, Cancer Incidence, Mean Difference in Serum Concentration between Cases and Controls (Std Error)						
Serum	**All**	**Lung**	**Breast**	**Prostate**	**GI**	

Concentration (µg/L)	**Cancer**					
	-0.6 (2.5)	7.4 (6.3)	5.4 (6.6)	1.7 (7.5)	-18.4 (7.7)	
No. Cases	111	17	14	11	11	
Connett et al., 1989, Cancer Deaths, Mean Difference in Serum Concentration between Cases and Controls						
Serum Concentration (µg/dL)	**All Cancer**	**Lung**				
	-1	-3.1				
No. Cases	156	66				

Table 8. Summary of Outcome Specificity Analysis of Studies Investigating the Association between Vitamin E/α-Tocopherol and Cancer

Exposure Variable	**Less Specific Outcome**	**More Specific Outcomes**				
Willett et al., 1984, Cancer Incidence, Mean Difference in Serum Concentration between Cases and Controls (Std Error)						
Serum Concentration (mg/dL)	**All Cancer**	**Lung**	**Breast**	**Prostate**	**GI**	
	-0.05 (0.06)	0.13 (0.15)	-0.16 (0.17)	-0.09 (0.19)	-0.15 (0.2)	
No. Cases	111	17	14	11	11	
Connett et al., 1989, Cancer Deaths, Mean Difference in Serum Concentration between Cases and Controls						
Serum Concentration (mg/dL)	**All Cancer**	**Lung**				
	0.03	-0.06				
No. Cases	156	66				

3. Comment on whether the method is general enough to be used directly, or if it can be extrapolated, for application to other chemicals and/or problem formulations. Please explain why or why not.

This method can be generalized to any chemical with enough epidemiological data available to test the above hypotheses of outcome severity, outcome specificity, and dose(exposure)-response. This method is restricted to human or ecological health questions.

4. Discuss the overall strengths and weaknesses of the method.

Strengths:

- Uses available epidemiology analyses
- Provides a more objective method that utilizes many epidemiology study results to assess causality
- Uses concepts that are grounded in statistics and toxicological theories to judge associations in epidemiology studies

Weaknesses:

- Relies on the available data and appropriate comparative analyses having been conducted
- Still not a definitive method for determining causality
- Assumes certain reliable patterns in the mathematical relationships and models
- Doesn't consider the internal validity of the epidemiology studies – we may need to add a quality assessment such as is used in systematic review methods

5. Outline the minimum data requirements and describe the types of data sets that are needed.

The minimum dataset is an epidemiology study that assesses multiple endpoint severities and specificities, as well as dose-response, for a relationship between a potential causative factor and a health endpoint. The types of datasets needed are epidemiology study methods and results. The preference would be to have multiple complementary studies for this analysis.

Does your case study:

A. Describe the dose-response relationship in the dose range relevant to human exposure?

Yes, the examples in the case study address dose-response in the range that is relevant to human exposure.

B. Address human variability and sensitive populations?

This case study does not specifically address sensitive populations, but it does address human variability with the considerations of outcome and exposure variance on the relationship between dose (exposure) and response.

C. Address background exposures or responses?

This case study does not address background exposures or responses.

D. Address incorporation of existing biological understanding of the likely mode of action?

As it currently stands, this case study does not incorporate existing biological understanding of the likely mode of action, but that is something that we would be very interested in discussing with the panelists. This case study does incorporate toxicological principles in the interpretation of epidemiology results.

E. Address other extrapolations, if relevant – insufficient data, including duration extrapolations, interspecies extrapolation?

This case study does not address other extrapolations, such as insufficient data, duration extrapolations, or interspecies extrapolation.

F. Address uncertainty?

This case study attempts to address uncertainty by evaluating many analyses of the same question (i.e. many epidemiology studies investigating the relationship between a particular exposure and response). This will help to address uncertainty by using data generated in different populations, with different co-exposures, etc. This case study also addresses the uncertainty around the assumptions of the effects of exposure measurement error and outcome measurement error, and whether it is valid to assume that a certain error will bias an effect estimate in a particular direction.

G. Allow the calculation of risk (probability of response for the endpoint of interest) in the exposed human population?

This case study does not directly allow the calculation of risk in an exposed human population, but it does help risk assessors decide whether it is appropriate to calculate risk based on the epidemiology data.

H. Work practically? If the method still requires development, how close is it to practical implementation?

The positive and negative controls provide evidence for the practical application of this case study. More work is required to test this idea before it is ready for practical implementation.

References

Alpha-Tocopherol, B.C.C.P.S.G., 1994. The effect of vitamin E and beta carotene on the incidence of lung cancer and other cancers in male smokers. N. Engl. J. Med. 330, 1029–35. https://doi.org/10.1056/NEJM199404143301501

Brakenhoff, T.B., van Smeden, M., Visseren, F.L.J., Groenwold, R.H.H., 2018. Random measurement error: Why worry? An example of cardiovascular risk factors. PLoS One 13, e0192298. https://doi.org/10.1371/journal.pone.0192298

Carrothers, T.J., Evans, J.S., 2000. Assessing the impact of differential measurement error on estimates of fine particle mortality. J. Air Waste Manag. Assoc. 50, 65–74.

Comstock, G.W., Helzlsouer, K.J., Bush, T.L., 1991. Prediagnostic serum levels of carotenoids and vitamin E as related to subsequent cancer in Washington County, Maryland. Am. J. Clin. Nutr. 53. https://doi.org/10.1093/ajcn/53.1.260S

Connett, J.E., Kuller, L.H., Kjelsberg, M.O., Polk, B.F., Collins, G., Rider, A., Hulley, S.B.,

1989. Relationship between carotenoids and cancer. The Multiple Risk Factor Intervention Trial (MRFIT) Study. Cancer 64, 126–34. https://doi.org/10.1002/1097-0142(19890701)64:1<126::aid-cncr2820640122>3.0.co;2-h

Freedman, N.D., Leitzmann, M.F., Hollenbeck, A.R., Schatzkin, A., Abnet, C.C., 2008. Cigarette smoking and subsequent risk of lung cancer in men and women: analysis of a prospective cohort study. Lancet. Oncol. 9, 649–56. https://doi.org/10.1016/S1470-2045(08)70154-2

Friedman, G.D., Blaner, W.S., Goodman, D.S., Vogelman, J.H., Brind, J.L., Hoover, R., Fireman, B.H., Orentreich, N., 1986. Serum retinol and retinol-binding protein levels do not predict subsequent lung cancer. Am. J. Epidemiol. 123, 781–789. https://doi.org/10.1093/oxfordjournals.aje.a114307

Fuller, W.A., 1987. Measurement Error Models. John Wiley & Sons, 1987.

Hausman, J., 2001. Mismeasured variables in econometric analysis: Problems from the right and problems from the left. J. Econ. Perspect. 15, 57–67. https://doi.org/10.1257/jep.15.4.57

Jurek, A.M., Greenland, S., Maldonado, G., 2008. How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null? Int. J. Epidemiol. 37, 382–5. https://doi.org/10.1093/ije/dym291

Jurek, A.M., Greenland, S., Maldonado, G., Church, T.R., 2005. Proper interpretation of non-differential misclassification effects: expectations vs observations. Int. J. Epidemiol. 34, 680–7. https://doi.org/10.1093/ije/dyi060

Klein, E.A., Thompson, I.M., Tangen, C.M., Crowley, J.J., Lucia, S., Goodman, P.J., Minasian, L.M., Ford, L.G., Parnes, H.L., Gaziano, J.M., Karp, D.D., Lieber, M.M., Walther, P.J., Klotz, L., Parsons, J.K., Chin, J.L., Darke, A.K., Lippman, S.M., Goodman, G.E., Meyskens, F.L., Baker, L.H., 2011. Vitamin E and the risk of prostate cancer: The selenium and vitamin E cancer prevention trial (SELECT). JAMA - J. Am. Med. Assoc. 306, 1549–1556. https://doi.org/10.1001/jama.2011.1437

Knekt, P., Aromaa, A., Maatela, J., Aaran, R.K., Nikkari, T., Hakama, M., Hakulinen, T., Peto, R., Teppo, L., 1990. Serum vitamin a and subsequent risk of cancer: Cancer incidence follow-up of the finnish mobile clinic health examination survey. Am. J. Epidemiol. 132, 857–870. https://doi.org/10.1093/oxfordjournals.aje.a115728

Lewer, D., McKee, M., Gasparrini, A., Reeves, A., de Oliveira, C., 2017. Socioeconomic position and mortality risk of smoking: evidence from the English Longitudinal Study of Ageing (ELSA). Eur. J. Public Health 27, 1068–1073. https://doi.org/10.1093/eurpub/ckx059

Loken, E., Gelman, A., 2017. Measurement error and the replication crisis The assumption that measurement error always reduces effect sizes is fals. Science (80-.). 355, 582–584. https://doi.org/10.1126/science.aam5409

Menkes, M.S., Comstock, G.W., Vuilleumier, J.P., Helsing, K.J., Rider, A.A., Brookmeyer, R.,

1986. Serum beta-carotene, vitamins A and E, selenium, and the risk of lung cancer. N. Engl. J. Med. 315, 1250–4. https://doi.org/10.1056/NEJM198611133152003

Nomura, A.M.Y., Stemmermann, G.N., Heilbrun, L.K., Salkeld, R.M., Vuilleumier, J.P., Nomura, A.M.Y., 1985. Serum Vitamin Levels and the Risk of Cancer of Specific Sites in Men of Japanese Ancestry in Hawaii. Cancer Res. 45, 2369–2372.

Omenn, G.S., Goodman, G.E., Thornquist, M.D., Balmes, J., Cullen, M.R., Glass, A., Keogh, J.P., Meyskens, F.L., Valanis, B., Williams, J.H., Barnhart, S., Cherniack, M.G., Brodkin, C.A., Hammar, S., 1996. Risk factors for lung cancer and for intervention effects in CARET, the beta-carotene and retinol efficacy trial. J. Natl. Cancer Inst. 88, 1550–1559. https://doi.org/10.1093/jnci/88.21.1550

Ordóñez-Mena, J.M., Schöttker, B., Mons, U., Jenab, M., Freisling, H., Bueno-de-Mesquita, B., O'Doherty, M.G., Scott, A., Kee, F., Stricker, B.H., Hofman, A., de Keyser, C.E., Ruiter, R., Söderberg, S., Jousilahti, P., Kuulasmaa, K., Freedman, N.D., Wilsgaard, T., de Groot, L.C., Kampman, E., Håkansson, N., Orsini, N., Wolk, A., Nilsson, L.M., Tjønneland, A., Pająk, A., Malyutina, S., Kubínová, R., Tamosiunas, A., Bobak, M., Katsoulis, M., Orfanos, P., Boffetta, P., Trichopoulou, A., Brenner, H., Consortium on Health and Ageing: Network of Cohorts in Europe and the United States (CHANCES), 2016. Quantification of the smoking-associated cancer risk with rate advancement periods: meta-analysis of individual participant data from cohorts of the CHANCES consortium. BMC Med. 14, 62. https://doi.org/10.1186/s12916-016-0607-5

Powell, H.A., Iyen-Omofoman, B., Hubbard, R.B., Baldwin, D.R., Tata, L.J., 2013. The association between smoking quantity and lung cancer in men and women. Chest 143, 123–129. https://doi.org/10.1378/chest.12-1068

Remen, T., Pintos, J., Abrahamowicz, M., Siemiatycki, J., 2018. Risk of lung cancer in relation to various metrics of smoking history: A case-control study in Montreal 11 Medical and Health Sciences 1117 Public Health and Health Services. BMC Cancer 18, 1–12. https://doi.org/10.1186/s12885-018-5144-5

Szpiro, A.A., Paciorek, C.J., Sheppard, L., 2011. Does more accurate exposure prediction necessarily improve health effect estimates? Epidemiology 22, 680–685. https://doi.org/10.1097/EDE.0b013e3182254cc6

Thun, M.J., Carter, B.D., Feskanich, D., Freedman, N.D., Prentice, R., Lopez, A.D., Hartge, P., Gapstur, S.M., 2013. 50-Year trends in smoking-related mortality in the United States. N. Engl. J. Med. 368, 351–364. https://doi.org/10.1056/NEJMsa1211127

Wald, N.J., Thompson, S.G., Densem, J.W., Boreham, J., Bailey, A., 1988. Serum beta-carotene and subsequent risk of cancer: Results from the BUPA study. Br. J. Cancer 57, 428–433. https://doi.org/10.1038/bjc.1988.97

Willett, W.C., Polk, B.F., Underwood, B.A., Stampfer, M.J., Pressel, S., Rosner, B., Taylor, J.O., Schneider, K., Hames, C.G., 1984. Relation of serum vitamins A and E and carotenoids to the risk of cancer. N. Engl. J. Med. 310, 430–4. https://doi.org/10.1056/NEJM198402163100705

Zeger, S.L., Thomas, D., Dominici, F., Samet, J.M., Schwartz, J., Dockery, D., Cohen, A., 2000. Exposure measurement error in time-series studies of air pollution: concepts and consequences. Environ. Health Perspect. 108, 419–26.