UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF COLUMBIA

UNITED STATES OF AMERICA,	)	
	)	
Plaintiff,	)	Civil Action No. 99-CV-02496 (GK)
	)	
v.	)	
	)	
PHILIP MORRIS USA INC.,	)	Next scheduled appearance:
f/k/a PHILIP MORRIS INC, *et al.*,	)	Trial (ongoing)
	)	
Defendants.	)	
	)	

WRITTEN DIRECT EXAMINATION OF
DONALD B. RUBIN, Ph.D.
SUBMITTED PURSUANT TO ORDER # 471

TABLE OF CONTENTS

I. INTRODUCTION.

A. GENERAL BACKGROUND.

Q. Would you please tell us your full name and where you live?

A. My name is Donald B. Rubin, Ph.D., and I live in Newton, Massachusetts.

Q. Please briefly describe your educational history.

A. For college, I attended Princeton University, where I earned a bachelor's degree with a major in psychology in 1965. I attended Harvard University for my post-graduate work. At Harvard, I earned a master's of science in computer science in 1966 and a Ph.D. in statistics in 1970.

Q. Did you receive any honors or awards while you were in school?

A. Yes, several, including membership in Phi Beta Kappa, the Woodrow Wilson Society and two National Science Foundation graduate fellowships.

Q. Where are you employed now?

A. Harvard University.

Q. What is your position at Harvard?

A. Since 2002, I have been the John L. Loeb Professor of Statistics in Harvard's Department of Statistics. From 1984 to 2002, I was a Professor in Harvard's Department of Statistics.

Q. Have you had any administrative responsibilities at Harvard?

A. Yes. I served three three-year terms as Chairman of Harvard's Department of Statistics, and, last year I completed a fourth term that was four years.

Q. Have you taught at other universities?

A. Yes, I was a Professor of Statistics at the University of Chicago before I came to Harvard. In addition, I have lectured on a visiting basis at several universities, including Princeton, the University of Texas, the University of Wisconsin, the University of Minnesota, and the

University of California at Berkeley. Those are all identified in my curriculum vitae (JD-068078) which was attached to my April 30, 2005 expert report in this case.

Q. Would you tell the Court about the primary subject areas of the courses you have taught at Harvard?

A. I have taught courses on causal inference, missing data methods, sample survey methods, Bayesian analysis and applied statistics in the Department of Statistics, the Department of Psychology, and the Department of Economics.

Q. Have you taught courses in which any other experts in this case were enrolled?

A. Yes, at Princeton in the 1970s, I taught a Ph.D.-level course on causal inference that one of the plaintiff's experts, Dr. Scott Zeger, took as a graduate student.

Q. What are your principal academic research interests?

A. For at least three decades, my research interests have focused primarily on causal inference and methods to address problems of missing data.

Q. Please describe very generally your work on causal inference?

A. The framework very briefly is as follows. To assess the causal effect of a past event (such as an alleged RICO violation of the defendants in 1970) on a particular outcome (such as smoking behavior in 1980), we compare the outcome in the world as it existed to what that outcome would have been at the same point in time but in a counterfactual world in which the event did not occur. Because that counterfactual world cannot be observed, we must use actual world data, together with explicit assumptions, to estimate what the outcome would have been in a counterfactual world in which the event did not occur. Similarly, to assess the causal effect of a future event on a particular outcome (*e.g.*, smoking prevalence) in the future, you compare the world as it would exist without the event to the world as it would with the event. Neither world

can be observed now, but we can use actual world data, together with explicit assumptions, to estimate what the outcomes will be in those two different scenarios. In this way, we can predict what would happen with and without the event, and, in turn, estimate the causal effect of the event.

Q. Has your general approach to causal inference been accepted by the scientific community?

A. Yes, I think so. Although I readily acknowledge that there are philosophical and statistical antecedents to my general approach, the literature sometimes refers to that approach as the "Rubin causal model" for my extensions and mathematical formulations of this idea. My articles on this topic frequently are cited in the literature, including the 2004 Surgeon General's Report (U.S. Exhibit 88,847) at 19, which essentially adopted an interpretation of cause that is consistent with my approach. Moreover, the testimony of Professor Heckman on this topic from the perspective of an economist is generally consistent with my approach.

Q. Have you received any awards for your academic research, particularly your work on causal inference and missing data?

A. Yes, I have received the Samuel S. Wilks Medal, which is the American Statistical Association's most prestigious award, for my lifetime contributions, including my work on causal inference and missing data, as well as the Parzen Prize for Statistical Innovation. My CV lists other awards I have received.

Q. Do you belong to any professional organizations or associations?

A. Yes. I have been elected a Fellow of the American Statistical Association, of the Institute of Mathematical Statistics, and of the American Association for the Advancement of Science. In addition, I have been elected as a member of the International Statistical Institute, the New York

Academy of Sciences, and the American Academy of Arts and Sciences. I also am a member of a variety of other professional associations, such as the American Public Health Association.

Q. Have you edited scientific journals?

A. Yes, both in my field of statistics and in related fields. I have been editor and/or associate editor of some of the most important peer-reviewed journals in statistics, including the *Journal of the American Statistical Association* and *Biometrika*. I also have served as an associate editor of journals in other fields, such as *Behavioral and Brain Sciences* and the *International Journal of Psychiatry in Medicine*.

Q. Have you served on advisory committees?

A. Yes, I have served on advisory committees for the National Academy of Sciences, professional associations (*e.g.*, the American Psychological Association, the American Statistical Association, and the Institute of Mathematical Statistics), and university departments (*e.g.*, Columbia University (statistics) and the University of California at Los Angeles (biostatistics and statistics)).

Q. Have you consulted with agencies or departments of the federal government?

A. Certainly. I have consulted on statistical issues with, among others, the Food and Drug Administration (the "FDA"), the Department of Health and Human Services' Centers for Disease Control and Prevention (the "CDC"), the General Accounting Office, the National Center for Health Statistics, the Internal Revenue Service, the Department of Defense, and the Census Bureau.

Q. Can you provide an example?

A. Sure. For the past several years, I have been consulting with the CDC in designing the data collection for, and planning the analysis of, its ongoing anthrax vaccine trials. That work, at

present, focuses on dealing with missing data issues. In brief, those trials seek to find a dosing regime for the vaccine that provides full protection to individuals against anthrax exposure with minimum negative side effects.

Q. Have you done consulting work for other entities?

A. Yes. I have consulted with pharmaceutical companies, often helping to design experiments and observational studies for submission to the FDA. For instance, I have been and am currently involved in a project for Genzyme in its submissions to the FDA for Farbrazyne, a synthetic enzyme designed to treat patients with Fabry's disease, as well as with Novartis, completing documentation for the use of Trileptal for mono therapy in children to control epileptic seizures, a submission that required "bridging" results from adults to children.

Q. Have you participated in applied statistical projects outside the pharmaceutical industry?

A. Sure. Recently, I have been involved in projects examining the causal effects of school choice programs on student performance, the economic effects of job training programs, and the causal effects of peer smoking on smoking initiation among college undergraduates.

Q. What is the rate you are charging for your work in this case?

A. I am charging $1,250 per hour for consulting and $1,600 per hour for testifying.

Q. Have you testified as a defense expert in other smoking and health actions?

A. Yes, I testified at trial once, in 1998 in *State of Minnesota v. Philip Morris Inc., et al.*, No. C1 94-8565 (2nd Dist., Ramsey County, Minn.). Also, between approximately 1997 and 2002 I was deposed in several cases where, in general, third-party plaintiffs – for example, states, unions, and insurers – sought reimbursement for smoking-related medical costs.

Q. Roughly what portion of your consulting income comes from your work with these defendants?

A. It has varied over time. Last year, work for these defendants amounted to between a quarter and a third of my gross consulting income.

II. OVERVIEW OF OPINIONS IN THIS CASE.

Q. Professor Rubin, before we discuss them in detail, would you give the Court a brief overview of your opinions?

A. Yes, my opinions are:

(1) There is a statistically reliable and valid way to estimate any possible causal effects of the defendants' alleged RICO violations on smoking behaviors. None of the experts for the United States whose testimony I have reviewed has even attempted to estimate, in a statistically valid and reliable manner, whether the defendants' alleged RICO violations caused increased smoking and, if so, the magnitude of any such increase.

(2) Defendants' alleged information-based RICO violations -- their alleged suppression and misrepresentations of the health risks of smoking, including addiction -- had essentially no causal effect on either adolescent smoking initiation or smoking cessation.

(3) There is a statistically reliable and valid manner to estimate the causal effects on smoking cessation from a national smoking cessation program, such as the one described by Dr. Fiore. Based on my review of Dr. Fiore's expert report, his deposition testimony, and his written direct examination, Dr. Fiore has not attempted to estimate reliably or validly the causal effects of such a national smoking cessation program.

(4) Dr. Wyant's estimates of deaths, treatment-years-of-disease, years of potential life lost, and health-care costs "attributable" to smoking are not statistically reliable and valid estimates of the causal effects of smoking.

III. THERE IS A STATISTICALLY RELIABLE AND VALID WAY TO ESTIMATE ANY POSSIBLE CAUSAL EFFECTS OF THE DEFENDANTS' ALLEGED RICO VIOLATIONS ON SMOKING BEHAVIORS.

A. GENERAL APPROACH.

Q. Very generally, what should be done to assess in a statistically reliable and valid way whether the Defendants' alleged RICO violations caused increased smoking?

A. In very general terms, evaluating any effects the alleged RICO violations had on smoking behavior involves comparing smoking behavior in the "actual" world, with the alleged RICO violations, to smoking behavior in a "counterfactual world" without the alleged RICO violations. The difference is the estimated causal effect of the alleged RICO violations on smoking behavior.

Q. Is it possible to estimate in a statistically reliable and valid way what smoking behavior would have been in the "counterfactual world" without the alleged RICO violations?

A. Yes. It is possible to estimate what smoking behavior would have been in the "counterfactual world" in which the alleged RICO violations did not occur, even though that "counterfactual world" does not exist and cannot be observed directly. To do that estimation correctly, we use available data from the "actual" world, as it exists, together with certain explicit assumptions, that permit us to "bridge" from the actual world data to the counterfactual world.

Q. What do you mean by assumptions that permit you to "bridge" from the actual to the counterfactual world?

A. It is commonplace in statistics to "bridge" results across subpopulations, time, and similar treatments. For example, "bridging" across time is extremely common -- it occurs whenever a

study conducted in the recent past is used to draw inferences about what will hold in the near future. Also, more extreme "bridging" commonly occurs.

That said, given the choice, one always would prefer to have data that directly address the question of interest. But, this is frequently impossible, particularly in estimating the "counterfactual world" in some past time. In these circumstances, one must "bridge" from the available data, using explicit assumptions; the only other alternative would be to throw up our hands and say we can draw no causal inferences, which would be scientifically nihilistic.

Q. Please provide the Court with an example of "bridging" from your own work.

A. Certainly.

As I mentioned earlier, I have consulted with the CDC concerning the design and analysis of studies to estimate the appropriate dosing regimes for anthrax vaccine. Ultimately, the goal is to estimate the amount of, and the manner for, delivering the vaccine necessary to protect people from an otherwise lethal dose of anthrax. Plainly, one cannot perform an experiment in which people, who have received varying amounts of the vaccine, including only a placebo, are exposed to a presumably lethal dose of anthrax.

Instead, the central idea and approach, simplified for our purposes here, was this: People and monkeys are given varying amounts of the vaccine. The monkeys are exposed to normally lethal doses of anthrax to determine the dosing regime that was appropriate for monkeys. To estimate doses for humans from the appropriate dosing regime for monkeys, studies were conducted examining the amount of vaccine which, when given to people, generated the same immunological response (*e.g.,* the same concentration of antibodies) in humans as in the surviving monkeys, assuming that humans with the same immunological response would also survive.

This is an example of "bridging" from studies of non-human primates to humans. It is discussed in one my peer-reviewed articles, Rubin, D.B., "Direct and Indirect Causal Effects Via Potential Outcomes," *The Scandinavian J. of Statistics,* 31:161-170 (2004) with discussion at 189-195 and reply at 196-198 (JD-068091).

Q. Were there data to address the propriety of "bridging" using immunological responses, such as antibody concentrations?

A. Yes. Specifically, studies of "passive vaccination" were also conducted, where blood with a specific concentration of antibodies was transfused from monkeys that had been given the vaccine into monkeys who had received no vaccine. The monkeys receiving the transfusion of blood with the antibodies were then exposed to anthrax to see if they, too, were protected.

Q. Who will decide whether the data from these studies are adequate for "bridging" results from monkeys to humans?

A. The FDA must approve both the vaccine and the dosing regimes.

Q. You said earlier that, given the choice, we would rather have data that directly address the issue of interest, but that those data may not be obtainable and that, in those instances, "bridging" based on the available data, together with explicit assumptions, may be used.

In the anthrax example you just discussed, what would have been the directly relevant data?

A. Data based on exposing humans to anthrax, which, obviously, we could not do.

Q. What, in the anthrax example, was the explicit assumption upon which the "bridging" from the non-human primates to humans was based?

A. The explicit assumption was that essentially equivalent immunological responses, such as concentrations of antibodies, meant essentially equivalent protection against anthrax in both species. Empirical support for that assumption is sought in various ways, including the passive vaccination studies and experiments with other species, such as guinea pigs.

Q. Are there any other examples from your own work that illustrate the principle of "bridging" with varying amounts of directly relevant data?

A. Yes, I have two examples that I think are helpful.

The first example concerns an analysis at Novartis, which had developed a drug to treat epilepsy in both adults and children, called Trileptal. The drug was approved for use as both an adjunctive, *i.e.*, in addition to other treatments, and a mono therapy, *i.e.*, as the only treatment, in adults based on randomized clinical trials.

Initially, Trileptal was approved for use in children *only* as an adjunctive therapy. Randomized clinical trials had been conducted with children on the use of the drug as an adjunctive therapy, but the ethics of conducting a randomized experiment using this drug as a mono therapy in children, in which some epileptic children would be randomly assigned to receive only a placebo, precluded this study.

FDA eventually approved the drug for use as a mono therapy among children, even though directly relevant data were unavailable -- randomized clinical trials of the drug as a mono therapy in children were never conducted. FDA approval was based on work to which I contributed that "bridged" the adult mono therapy results in a statistically appropriate way to the children, using the adjunctive therapy results from both adults and children. *See* NDA# 21-014/S-003.

The second example concerns alternative treatments for breast cancer -- lumpectomy and mastectomy. In six major randomized clinical trials in the 1980s, lumpectomy had been shown to be as efficacious as mastectomy among certain groups of women. However, the women and doctors who participated in the randomized trials (a) were willing to have this important and personal treatment choice be assigned randomly; and (b) were treated at highly specialized medical research centers that were willing to participate in randomized trials and collect all appropriate data for years.

The General Accounting Office was interested in assessing empirically whether the results of the randomized clinical trials -- that lumpectomy was equally efficacious -- could be extrapolated to the general population of breast cancer patients in the U.S. with the same conditions. Working with the GAO, I helped design an observational study to assess this question. The observational study found that lumpectomy appeared equally efficacious even among women who chose, with their doctors, the form of surgery and who were not treated in highly specialized research centers. This work is briefly described in my peer-reviewed article, "Estimating Causal Effects From Large Data Sets Using Propensity Scores," *Annals of Internal Medicine*, 127(8, part 2):757-763 (1997) (JD-063884), which includes references to the underlying GAO report.

Q. Is "bridging" ever completely avoidable?

A. No, it is virtually impossible to eliminate "bridging" altogether. "Bridging" occurs whenever the results of a study are extrapolated beyond the study itself, including extrapolations to the future or to other individuals who were not part of the study.

For example, the observational study of lumpectomy versus mastectomy "bridged" in time from women who had breast cancer at the time the observational study was conducted to

women who were diagnosed with breast cancer later in time. It also "bridged" from the study participants to those who did not participate in the study -- for example, other women who were contemporaneously diagnosed with breast cancer and other doctors who contemporaneously treated breast cancer patients.

Q. Do plaintiffs' experts employ bridging?

A. Yes. For example, Dr. Fiore's reports (Fiore et al 1996, 2000) and the reports from the Cochrane Collaboration on which he relies summarize results from hundreds of experiments in the past and in various countries to infer causal effects of various smoking cessation interventions in future United States subpopulations.

Q. Is your approach to assessing the causal effect of Defendants' alleged misconduct a topic that you have addressed in your peer-reviewed publications?

A. Yes. In addition to my work on causal inference generally, I have specifically discussed in my peer-reviewed publications how to estimate the causal effects of the tobacco industry's alleged misconduct on smoking behavior and, if appropriate, other endpoints, such as health-care costs.

For example, this issue was the focus of my chapter, "Statistical Issues in the Estimation of the Causal Effects of Smoking Due to the Conduct of the Tobacco Industry," in *Statistical Science in the Courtroom*, J. Gastwirth (ed.), New York: Springer-Verlag, pp. 321-351 (2000) (JD-063877); as well as my more technical article, "Estimating The Causal Effects of Smoking," in *Statistics in Medicine,* 20:1395-1414 (2001) (JD-063879). I also discuss more specific aspects of a correct design in my article "Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation," *Health Services & Outcomes Research Methodology,* 2:169-188 (2001) (JD-063885).

Q. Is that a topic that you have also addressed specifically in invited lectures?

A. Yes. Estimating the causal effects of the tobacco industry's alleged misconduct has been a topic I have addressed in invited lectures, including two invited lectures to the Centers for Disease Control and Prevention; invited lectures at various universities, including Columbia and the University of Pennsylvania; as well as in my invited 2001 Lowell Reed Lecture to the American Public Health Association.

B. SPECIFIC STEPS IN ESTIMATING THE CAUSAL EFFECTS ON SMOKING BEHAVIOR OF THE TOBACCO INDUSTRY'S ALLEGED MISCONDUCT.

Q. What principal features must an analysis possess in order to estimate in a reliable and valid way the causal effect of the Defendants' alleged misconduct?

A. In summary form, an analysis that attempts to estimate the causal effect of the Defendants' alleged misconduct must have a number of principal features, including the following:

(1) It must define the population of individuals for which we want to generate an estimate. For example, we might be interested in assessing causal effects for the entire U.S. population or we might want to focus on a subset of individuals, such as Dr. Gruber's so-called "youth-addicted population."

(2) It must define the specific acts of the Defendants that are alleged to have violated RICO in each year, so that we can define the counterfactual world without those acts.

(3) It must separately analyze effects in defined, relevant subpopulations.

(4) It must estimate, using data and explicit assumptions, the smoking behavior of the subpopulations in the actual world.

(5) It must estimate, using data and explicit assumptions, the smoking behavior of the subpopulations in the counterfactual world, without the alleged misconduct.

(6) The model must compare the outcomes generated in steps (4) and (5) in different subpopulations and then combine them across subpopulations.

The differences in the estimates of smoking behavior in the two worlds -- the actual (step 4) and counterfactual (step 5) -- generates, if the other steps were done correctly, a statistically valid and reliable estimate of the causal effect of the alleged misconduct on smoking behavior.

Of course, if we also wanted to know the effect of any change in smoking behavior on particular endpoints, such as health care costs, then we would need to incorporate a separate model relating smoking behavior to health-care costs in steps 4 and 5. Then, step 6 would end up comparing health-care expenditures in the actual world to health-care expenditures in a counterfactual world without the defendants' alleged misconduct.

Q. What do you mean by "defined, relevant subpopulations?"

A. Consider first the effect of defendants' alleged misconduct on smoking behavior. Then relevant subpopulations are subgroups of the total population that could have responded differently to the alleged misconduct.

For instance, a particular action in 1970 that allegedly appealed to youth presumably could have had a different effect on the smoking behaviors of individuals who were youth in 1970 than on the smoking behaviors of individuals who were retired in 1970.

More generally, just as Surgeon General Carmona testified in his written direct (at 16-17) that smoking cessation initiatives must be "culturally competent and targeted" toward specific "racial/ethnic populations," so, too, it is reasonable to anticipate that any effect of the Defendants' allegedly fraudulent statements could have been different in different subpopulations. *See also* Surgeon General Carmona's *U.S.* Trial Tr. 5/3/05 a.m. at 20107-20108 and 20160. Indeed, the literature that evaluates the efficacy of different interventions designed to

reduce smoking initiation suggests that different subpopulations of adolescents respond differently to different messages. *See, e.g.,* NCI Monograph 14, *Changing Adolescent Smoking Prevalence,* (2002) (Burns, D., senior scientific ed.) (U.S. Exhibit 72,977), which identifies educational aspirations as critical to smoking behavior; explores the effect of race on smoking behavior; and characterizes self-described school performance as "[o]ne of the most powerful predictors of adolescent smoking behavior" (at 85).

Q. Is it important to estimate effects separately within subpopulations?

A. Absolutely. A hypothetical example illustrates this concept. In the statistical literature extreme examples such as this are referred to as "Simpson's paradox." The four cells in the table below display a purely hypothetical situation where (a) male smokers and male never-smokers have identical health-care expenditures, and (b) female smokers and female never-smokers have identical health-care expenditures. Thus, for neither males nor females is there any evidence that smoking is associated with increased expenditures. Nevertheless, the average expenditures for all smokers are 50% larger than those for never smokers. To get the correct causal inference in this hypothetical illustration, one must control for gender because it is associated with both smoking (males tend to smoke more) and expenditures (males tend to have more expenditures).

Illustration of Simpson's Paradox

	Smokers	Never-Smokers
Males	*n*=200 ave=$2000	*n*=50 ave=$2000
Females	*n*=50 ave=$1000	*n*=200 ave=$1000
All	*n*=250 ave=$1800	*n*=250 ave=$1200

Q. What specific lesson can be drawn from this illustration?

A. This illustration demonstrates that if we were to compare equal sample sizes of smokers and non-smokers, we would come to the conclusion that smokers have higher health care costs than do non-smokers. An accurate assessment of this hypothetical example is that there is no difference between the health care costs of smokers and non-smokers, either male or female.

Q. Have any of the Government's experts whose reports and testimony you have reviewed provided the Court with a statistically valid and reliable estimate of the effect of the Defendants' alleged RICO violations on smoking behavior?

A. No. I have reviewed reports and testimony from Drs. Biglan, Burns, Eriksen, Fisher, Fiore, Gruber, Miller, Wyant, and Zeger. I have not seen any quantitative estimate, much less one that was statistically valid and reliable, of the effect of the Defendants' alleged RICO violations on smoking behavior.

IV. DEFENDANTS' ALLEGED INFORMATION-BASED RICO VIOLATIONS -- THEIR ALLEGED SUPPRESSION AND MISREPRESENTATIONS OF THE HEALTH RISKS OF SMOKING, INCLUDING ADDICTION -- HAD ESSENTIALLY NO CAUSAL EFFECT ON ADOLESCENT INITIATION OR SMOKING CESSATION.

A. ADOLESCENT SMOKING INITIATION.

Q. Professor Rubin, can one apply general principles of causal inference to address the question: Did more youth start smoking because of Defendants' alleged information-based fraud concerning diseases and conditions caused by smoking, and, if so, how many?

A. Yes. To address that causal question validly and reliably, one must apply sound statistical principles of causal inference.

Q. Did you attempt to do that in this case?

A. Yes.

Q. What did you do?

A. Against the backdrop of the framework for estimating the effects of the tobacco industry's alleged misconduct that is in my peer-reviewed publications, I systematically reviewed the available empirical evidence in the literature on efforts to prevent youth smoking initiation, focusing particularly on studies that examined interventions that first provided youth with more factual information about the disease and conditions caused by smoking, including its addictive nature, and then assessed the causal effects of providing that information.

Q. Is there a substantial body of literature reporting empirical evidence on the effect of providing adolescents with more factual information about the health risks of smoking on their subsequent smoking initiation?

A. The literature on interventions to curtail youth smoking is voluminous. I reviewed every published article I could find -- over 600 in all, identified in JD-068089. Despite my effort to be

comprehensive, it is certainly possible that there are additional studies that I and my research assistant could not find.

Q. Did you limit your review to studies that presented new data?

A. No. Although most of the articles I reviewed presented new data, I also examined review articles that summarized the literature, including meta-analyses and pertinent portions of various Surgeons General Reports.

Q. Was every article that you reviewed pertinent to the question of whether providing adolescents with additional information on the health risks of smoking prevented smoking initiation?

A. No, many were not.

It was important for this purpose of mine that the underlying study examined the effect of providing adolescents with additional information on the health risks of smoking. A study was pertinent for my purpose if a major component of the intervention consisted of providing adolescents with additional health-related information, even if the intervention had other components.

On the other hand, I did not consider pertinent for this part of my work studies that examined the effect on adolescent smoking initiation of other interventions that either (a) had no informational component, but, rather, focused on other intervention features, such as heightened restrictions on public smoking or better enforcement of local laws prohibiting the sale of cigarettes to minors; (b) had only a minor informational component; or (c) had nearly identical informational components in both the treatment and control groups.

Q. When the intervention did more than provide adolescents with additional information about the health risks of smoking, were you able to isolate the effect of providing additional information alone?

A. It is possible to design a study that would permit one to isolate the separate effect of each component of an intervention. Even if a study is not designed in that fashion, it may be possible in some circumstances, with the raw data, to attempt to isolate statistically the separate effects of different components of the intervention, but this is not easy to do well.

Even when the effect of providing additional health-risk information alone is not separately estimated, those studies are still pertinent because at least one major part of the intervention included an informational component. For those studies, any causal effects are for the intervention as a whole. Assuming that no component of the intervention increased the likelihood of smoking initiation among adolescents, the studies that did more than simply provide adolescents with incremental additional information on the health risks of smoking may be interpreted as estimating the upper bound of the causal effect of providing additional health risk information alone.

Q. In forming your opinions in this case, did you give equal weight to each of the pertinent studies in the literature?

A. No.

Q. Why not?

A. First, there was considerable variation in the statistical reliability and validity of the studies in the literature that examined the effect on adolescent smoking initiation of providing additional information about the health risks of smoking.

Some of the studies were excellent. They were randomized experiments, they were well designed from a statistical perspective, and they were analyzed in a statistically appropriate way. Other studies fell well short of this ideal.

In forming my opinions, I first gave greater weight to those studies that were statistically reliable and valid.

Q. What were the features or characteristics that distinguished the statistically reliable and valid studies from the others?

A. There were a number of factors I considered. For example:

I had a strong preference for randomized experiments as opposed to observational studies. Nobody seriously disputes that data from well-conducted experiments provide a more reliable basis for causal inference for the subpopulations included in the experiment. *See, e.g.*, Federal Judicial Center's *Reference Manual on Scientific Evidence*, 2nd Ed. (2000) (JD-063827) at 93 ("Consequently, inferences based on well-executed randomized experiments are more secure than inferences based on observational studies.").

Among the experiments, the more relevant studies had long-term follow-up. Long-term follow-up is important to ensure that the intervention doesn't merely postpone the onset of smoking for a short period as opposed to preventing it. There is some evidence that certain interventions, for example, reduce experimentation with cigarettes in junior high school, but they appear to have no long-term effect on smoking rates through high school. *See, e.g,,* Bell et al., "Do Drug Prevention Effects Persist into High School? How Project ALERT Did with Ninth Graders," *Prev. Med.,* 22:463-483 (1993) (JD-064062); Flay et al., "Six-Year Follow-up of the First Waterloo School Smoking Prevention Trial," *Am. J. Pub. Health*, 79(10):1371-1376 (1989) (JD-064033).

From a statistical perspective, the better experimental studies also had low dropout rates. A substantial dropout rate complicates any conclusions. Typically, in a smoking curtailment study, dropout rates are greater in the group that receives the intervention, and the dropouts tend to be those who do not follow the advice of the intervention (*i.e.,* they tend to be smokers).

This problem has long been recognized in statistics generally and in the literature concerning smoking prevention and cessation specifically. For instance, Peterson et al., "Hutchinson Smoking Prevention Project: Long-term Randomized Trial in School-based Tobacco Use Prevention -- Results on Smoking," *J. Nat'l Cancer Inst.*, 92(24):1979-1991 (2000) (JD-062758) observed at 1982:

> It is well known that outmigrators are different from non-outmigrators -- e.g., with regard to smoking prevalence -- and, thus, must be included in the follow-up to ensure scientific integrity.

Plaintiff's experts in this case have also emphasized this issue in their own writings. Biglan et al., "Do Smoking Prevention Programs Really Work? Attrition and the Internal and External Validity of an Evaluation of a Refusal Skills Training Program," *J. of Behav. Med.,* 10(2):159-171 (1987) (JD-063805) noted (at 159, abstract) that the

> result may have been due to a higher rate of attrition among high-rate smokers in the treatment condition than in the control condition. Attrition also affected external validity. Across both conditions, subjects who were smoking at pretest and who were at risk to smoke were more likely to be missing at follow-up.

Fiore et al., *Smoking Cessation. Clinical Practice Guideline No. 18,* USDHHS, Rockville, MD: Public Health Service, Agency for Health Care Policy and Research, AHCPR Pub. No. 96-0692 (April 1996) (JD-063828) at 11 employed a modified "intent-to-treat analysis" to account for disproportionate dropout rates in smoking cessation interventions, where all dropouts were considered failures.

The data, of course, should be analyzed in a statistically appropriate way. For example, if entire schools were randomly assigned to receive an intervention, then the effect of the intervention must be assessed statistically at the school, as opposed to the individual student, level. That is, one must use the unit of analysis that conforms to the study design. If the wrong unit of analysis is used, statistically significant results may be found where none exists.

Jerome Cornfield, a contributor to the 1964 Surgeon General's Advisory Committee, wrote on this very point: "randomization by cluster accompanied by an analysis appropriate to randomization by individual is an exercise in self-deception, however, and should be discouraged." Cornfield, J., "Randomization by Group: A Formal Analysis," *Am. J. Epidem.,* 108(2):100-102 (1978) (JD-063816), at 101-102.

Q. You noted your strong preference for experiments. Are you suggesting that non-experimental or observational data cannot serve as the basis for causal inference?

A. No. It is more difficult, however, as recent experience with hormone replacement therapy shows. Namely, observational studies, such as those utilizing the nurses health study, had suggested that hormone replacement therapy reduced certain disease risks, but, in a subsequent large randomized experiment, the Women's Health Initiative, the therapy was shown to increase those disease risks. Of course, when using observational studies to draw conclusions, one must make sure that the observational studies are well-conducted.

Q. Is it possible to distinguish the well-conducted observational studies from the others?

A. Yes, it is.

Q. What features make for a well-conducted observational study?

A. There are five main features that distinguish an observational study that provides a sound basis upon which to draw a causal inference from those that do not. Those studies that have those five features I consider "well-conducted."

A well-conducted observational study is one that: (1) controls well for background differences, between those exposed to the intervention and those not so exposed, that might influence the outcome; (2) has enough independent units in both the exposed and control groups to provide reliable results; (3) has a study design and analysis that was chosen *a priori*; (4) minimizes attrition, dropout, and other forms of missing data, and when they occur, deals with them properly; and (5) has an analysis appropriate for the chosen design, as in good randomized experiments.

Q. Let's talk a little bit about some of those features to the extent that we haven't already addressed them in discussing the randomized experiments. We have already discussed items (4) and (5) from your list, now let's discuss the others. Why is it important to control for background variables?

A. In a randomized experiment, control for background variables is accomplished by randomly assigning a large enough initial group of participants to either receive a treatment (exposed) or not (not exposed). With random assignment of a large enough group of participants, one can safely assume that the two resulting groups -- treatment and control -- are essentially identical at the start of the study. With that approach, it is possible to attribute the differences between the groups to the intervention or treatment.

With an observational study, one must collect data on background characteristics and create treatment and control populations with approximately the same distribution of all relevant background characteristics. This can be accomplished either by creating subgroups of treatment

and control individuals with similar distributions of these background characteristics or by creating treatment-control matched pairs of people who have similar background characteristics. It is possible to use statistical modeling techniques to control for minor differences in background characteristics, although relying solely on such techniques has been known to be notoriously unreliable for decades. (Cochran and Rubin, "Controlling Bias in Observational Studies: A Review," *Sankhya*, A, 35(4):417-446 (1973) (JD-040397).)

Q. Why is it important to have enough independent exposed and control units?

A. Without a sufficient number of independent exposed and control units, the study will not be statistically reliable. Even a randomized experiment using only two schools -- one a control school and one a treatment school -- uses too small a number of units to be reliable because the school must be used as the appropriate unit, not the individual students.

Q. Please explain why it is important that the design and analysis be selected *a priori*?

A. It is important to select the design and analysis *a priori* so that they are disconnected from any observation of the outcomes to avoid potential manipulation of the resulting answers. In a randomized experiment, which one is trying to emulate in the construction of a well-conducted observational study, one does not have the opportunity to observe the data before selecting the design. The outcome data turn out to be what they will. In FDA drug trials, for instance, the one primary analysis of outcome data must be selected at the same time as the design for the study is chosen; one is not allowed to conduct repeated analyses hunting for a result favorable to the drug company.

In contrast, the data from observational studies are too often analyzed over and over, yielding inconsistent results from which the investigator can pick and choose. As Dr. Fisher, one

of the experts for the Government, wrote in "Statisticians, Econometricians, and Adversary Proceedings, *J. Amer. Stat. Ass'n*, 31(394):277-286 (1986) (JD-063831) at page 279:

> [T]he statistician cannot simply mine the data until a satisfactory model is found to fit the sample. To do so at best vitiates any claim as to the statistical properties of the model and at worst exposes the statistician to the suspicion that models have been rejected until one favoring the client is found.

It is possible to avoid this problem. Analyses of observational studies can be designed without access to any outcome data so that the design and analysis cannot be revised to yield a result that the investigator perceives is desirable. That is a hallmark of a well-conducted observational study. As I wrote in "Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation," *Health Services & Outcomes Research Methodology,* 2:169-188 (2001) (JD-063885) at 169:

> Arguably, the most important feature of experiments is that we must decide on the way data will be collected before observing the outcome data. If we could try hundreds of designs and for each see the resultant answer, we could capitalize on random variation in answers and choose the design that generated the answer we wanted! The lack of availability of outcome data when designing experiments is a tremendous stimulus for "honesty" in experiments and can be in well-designed observational studies as well.

Q. Are there examples of observational studies conducted with this characteristic?

A. Yes. One example is the observational study of the effectiveness of mastectomy versus breast conservation surgery in general practice, which I discussed earlier and which is summarized in Rubin, D.B., "Estimating Causal Effects From Large Data Sets Using Propensity Scores," *Annals of Internal Medicine*, 127 (8, part 2):757-763 (1997) (JD-063884). That study demonstrated that an observational study, if designed carefully and objectively, can achieve results consistent with those of randomized experiments.

Q. Are there any other issues that are important?

A. Yes. A final caution concerns the method of reporting results: one must be sure that the analyses are presented in a straightforward fashion so that they are not misunderstood.

Q. Were there instances in which you had to consider some studies that fell short of today's standards of statistical reliability and validity?

A. Yes. Fewer pertinent articles were published in the 1960s and early 1970s that meet today's standards for statistical reliability and validity. Although randomized experimental studies were conducted in the 1960s and early 1970s, the researchers were less sensitive to issues concerning dropout rates, the appropriate unit of analysis, and even the utility of randomization. Moreover, they communicated in their publications less information about the underlying data than the authors of many later articles.

1. DOES PROVIDING ADDITIONAL TRUTHFUL INFORMATION ABOUT THE HEALTH RISKS OF SMOKING, ABOVE AND BEYOND THAT WHICH WAS ALREADY AVAILABLE, AFFECT ADOLESCENT SMOKING INITIATION?

a. THE "INFORMATION DEFICIT" MODEL

Q. Were there interventions in the 1960s and 1970s that provided adolescents with additional truthful information about the health risks of smoking, above and beyond that which was readily available, in an effort to prevent smoking initiation?

A. Yes, there were a number of such studies in the 1960s and early 1970s. The interventions were based on a model of filling an "information deficit" about the long-term health risks of smoking, typically with messages designed to "arouse fear," as was noted in the 1994 Surgeon General's Report, *Preventing Tobacco Use Among Young People* (U.S. Exhibit 64,693), at 216-217:

> In the 1960s and early 1970s, strategies to prevent the onset of cigarette smoking were often based on the premise that adolescents who engaged in smoking behavior had failed to comprehend the Surgeon General's warnings on the health hazards

> of smoking. The assumption was that these young people had a deficit of information that could be addressed by presenting them with health messages in a manner that caught their attention and provided them with sufficient justification not to smoke.
>
> * * * *
>
> Through images and messages often intended to arouse fear, these programs were designed to convince the adolescent audience that persons who smoke risk a variety of serious physical consequences throughout their lives, including an increased likelihood of premature death in adulthood from cardiovascular disease or cancer.

Q. Was the "information deficit" approach successful in preventing smoking initiation among adolescents?

A. No. It was "not effective," as the 1994 Surgeon General's Report (U.S. Exhibit 64,693), at 217 noted:

> The underlying assumption of these information-focused programs proved to have limited grounding. Although expanded educational efforts in schools throughout the 1970s provided adolescents with various kinds of smoking-related information, this information alone did not deter them from beginning to smoke. Comprehensive reviews published at that time concluded that smoking-prevention programs based on the information deficit approach were not effective.

Q. Did any of the studies that employed the information deficit approach in the 1960s and 1970s report a statistically significant reduction in smoking?

A. Yes, but there was only one that I was able to find. It was a non-randomized study by Horn in 1960, "Modifying Smoking Habits in High School Students," *Children* 7(2):63-65 (1960) (JD-063847). Horn reported a statistically significant reduction in smoking initiation among the high school students who received information about smoking's long-term health risks, which he termed the "remote" intervention. The term "remote" refers to long-term health consequences, such as the increased risk of lung cancer, that would arise decades later, after an extended period of smoking.

However, apart from not being a randomized experiment, the data were misanalyzed, focusing on the individual student, not the school, as the unit of analysis, even though entire schools were assigned to receive or not receive the intervention. The effect of that error is to increase the likelihood of finding a statistically significant effect where none exists. The article, moreover, only incompletely reports the results.

The maximum effect reported in the 1960 Horn article is a 4 percent reduction in smoking after 8 months, but this result was not replicated shortly thereafter in better-conducted and better-reported studies. For instance, Beckerman, S.C., "Report of an Educational Program Regarding Cigarette Smoking Among High School Students," *J. Maine Med. Ass'n,* 5:60-71 (1963) (JD-063802), was a randomized experiment involving twenty-six schools. The intervention was modeled on that used by Horn and included the "remote" intervention. Even though the analysis erroneously was conducted at the student level, instead of the school, which, again, would increase inappropriately the likelihood of reporting a statistically significant result, there was no statistically significant difference in the rise in smoking between the schools that received the "remote" smoking prevention intervention and those that did not.

Likewise, Creswell, W.H. et al., "University of Illinois Anti-Smoking Education Study," *Illinois J. Educ.,* 60(3):27-37 (1969) (JD-063817), reported a relatively early study in which five high schools and twelve junior high schools were randomly assigned to receive a smoking prevention intervention including Horn's "remote" intervention. The analysis inappropriately was conducted at the student level instead of the school level and attrition from the study was not addressed. Both of these errors would tend to lead to an erroneous finding of statistical significance when, in fact, statistical significance is not achieved. Nonetheless, the study reports

no statistically significant difference in smoking initiation among students receiving the "remote" intervention.

Q. Are there studies that examine whether the "information deficit" approach in the 1960s and 1970s was "not effective" because the intervention failed to change students' beliefs and attitudes toward smoking?

A. Yes. Some of the studies in the 1960s that were better-conducted from a statistical perspective specifically examined this question of the effect of the information-based intervention on students' knowledge or attitudes toward smoking. They reported that adolescents assigned to receive the information-focused intervention reported statistically significant increases in knowledge and anti-smoking attitudes, but the change in knowledge and attitudes did not translate into changes in smoking initiation.

Q. To which studies do you refer?

A. Two from the 1960s.

First, in 1963, Beckerman (JD-063802) reported that students who received additional information about the health risks of smoking were more likely than the control group to agree at follow-up with anti-smoking statements, such as "smoking costs more than the pleasure is worth," or "when I have children, I hope they never smoke."

Second, in 1965, Monk, M. et al, "Evaluation of an Antismoking Program Among High School Students," *Am. J. Public Health,* 55(7):994-1004 (1965) (JD-064002), reported that students who received the information-based intervention were statistically significantly more likely to agree, at follow-up, that "smoking is dangerous to health" than the control group that did not receive the additional information.

In neither study did the intervention that provided more information about the health risks of smoking appear to prevent smoking initiation.

Q. Are you alone in concluding that the "information deficit" approach from the 1960s and early 1970s may have been successful in changing knowledge or attitudes toward smoking but that those changes did not translate into changes in smoking initiation among adolescents?

A. Not at all. It is often repeated in the literature.

For example, Flay, B.R., "Psychosocial Approaches to Smoking Prevention: A Review of Findings," *Health Psychology* 4(5):449-488 (1985) (JD-061642), reviewed these studies and noted (at 450):

> Most past programs have been based on the premise that if children know why cigarette smoking is bad for them, they should choose to not start smoking. Of those conventional smoking education programs evaluated, many have succeeded in changing students' knowledge, some their beliefs, and some their attitudes, but very few have consistently reduced the onset of smoking behavior.
>
> The failure of informational or fear programs to change behavior comes as no surprise to psychologists. (Citations omitted.)

Earlier, Greenberg et al., "Smoking Intervention: Comparing Three Methods in a High School Setting," *J. School Health*, 48(8):498-502 (1978) (JD-065614) noted (at 498):

> It would seem then that although public education has created an awareness of the risk of smoking, it has produced little, if any, significant effect on the general smoking habit of students. Unfortunately, few health education programs dealing with smoking reduction present any concrete suggestions or positive demonstrations on how to stop smoking.

And, Telch et al., "Long-Term Follow-Up of a Pilot Project on Smoking Prevention with Adolescents," *J. Behav. Med.*, 5(1):1-8 (1982) (JD-065904) noted (at 2):

> Numerous antismoking programs have been implemented in junior and senior high schools in attempts to dissuade adolescents from smoking. Traditionally, programs have employed a wide range of techniques including lectures, discussions, posters, and films aimed at increasing student awareness of the harmful long-term effects of cigarette smoking. While some studies have reported positive changes in knowledge and attitudes, most show little or no effect on students' reported smoking behavior (Andrus, 1964; Beckerman, 1963; Evans and Borgatta, 1970; Holland, 1968; Irwin *et al*., 1970; Jeffreys and Westaway, 1961; Morrison, 1964).

b. AFFECTIVE EDUCATION MODEL.

Q. What was the next major approach to the prevention of smoking initiation reported by the 1994 Surgeon General's Report?

A. It was the so-called "affective education" model. It was described in the 1994 Surgeon General's Report (U.S. Exhibit 64,693 at 217) as an attempt "to increase adolescents' perceptions of self-worth and to establish or clarify a health-related value system that would support a young person's decision not to smoke."

Q. Did interventions implementing the affective education model have an information-based component?

A. Many of them did, yes.

For instance, Schaps, E. et al., "Primary Prevention Research: A Preliminary Review of Program Outcome Studies," *Int'l J. Addictions*, 15(5):657-676 (1980) (JD-065835) reviewed 35 drug abuse prevention programs, many of which were "new generation" affective education programs and had an information-based component. *See also* Tobler et al., "School-Based Adolescent Drug Prevention Programs: 1998 Meta-Analysis," *J. Primary Prevention* 20(4):275-336 (2000) (JD-063898), at 301 and Table 9 (grouping studies into categories, including "Knowledge + Affective" group); Kinder, B.M. et al., "Drug and Alcohol Education Programs: A Review of Outcome Studies," *Int'l J. Addictions,* 15(7):1035-1054 (1980) (JD-

065662) (review of studies with description of intervention components); Hansen, W.B. et al., "Evaluation of a Tobacco and Alcohol Abuse Prevention Curriculum for Adolescents," *Health Educ. Quarterly,* 15(1):93-114 (1988) (same) (JD-064037).

Q. Did interventions that implemented the affective education model, which often had an information-based component, successfully prevent youth smoking initiation?

A. The statistically reliable and valid studies generally did not show a long-term decrease in smoking initiation in the schools that received the intervention. As the 1994 Surgeon General's Report (U.S. Exhibit 64,693) noted, at 217:

> Reviews based on more than a decade of research have concluded that interventions based on the affective education model were no more effective in reducing adolescent smoking than those based on the information deficit model. Some studies even suggested (that is, without conclusive findings) that these programs may have had the untoward effect of eliciting interest in the behaviors they attempted to discourage.

Likewise, Tobler et al., "School-Based Adolescent Drug Prevention Programs: 1998 Meta-Analysis," *J. Primary Prevention* 20(4):275-336 (2000) (JD-063898), at 302, concluded that "none of the Non-interactive groups in the set of high-quality evaluations show effects significantly different from zero." This general lack of effect for both the "Knowledge Only" and "Knowledge + Affective" interventions is vividly illustrated in Figure 4 (at 311) and reproduced below, which shows that the results from the 69 non-interactive substance abuse interventions, including those designed to curtail smoking, appear to bounce randomly above and below the horizontal line indicating zero estimated effect, with the biggest absolute effect being negative.

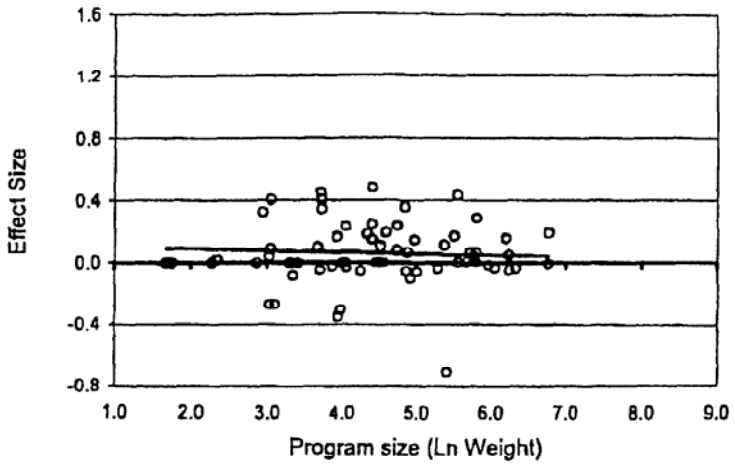


Fig. 4. 69 Non-interactive programs (fitted line is WLS regression).

The non-interactive interventions are more pertinent for assessing possible effects of the Defendants' alleged information-based misconduct concerning the health risks of smoking than the interactive interventions, which "provide contact and communication opportunities for the exchange of ideas among participants and encourage learning drug refusal skills." (JD-063898) at 287. Nonetheless, the effectiveness of these interactive interventions diminished when the programs increased in size. Tobler et al. (JD-063898) at 315 note:

> As the size approaches a few thousand students, the strength of the Interactive approach to prevention dwindles to match the lower level of the Non-interactive approach . . . .

The point is illustrated graphically by Tobler et al. in Figure 5 at 311, shown below, which makes it clear that the effectiveness of the smaller interactive approaches was not replicated as the programs increased in size:

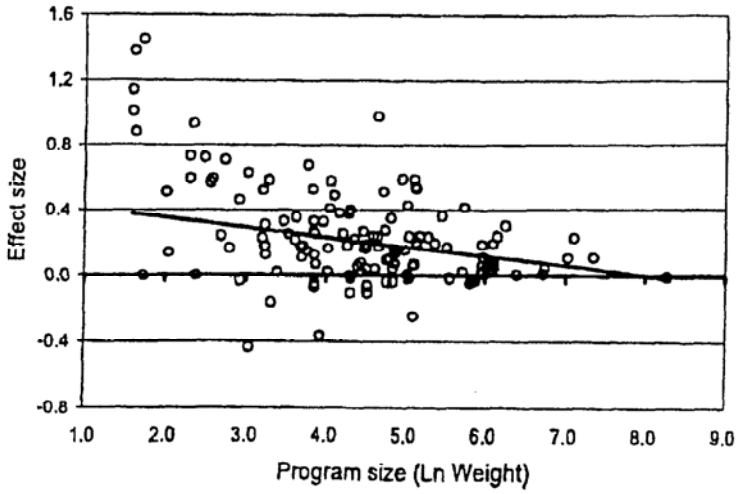


Fig. 5. 138 Interactive programs (fitted line is WLS regression).

Additionally, Botvin et al., "A Cognitive-Behavioral Approach to Substance Abuse Prevention: One-Year Follow-Up," *Addictive Behav.*, 15:47-63 (1990) (JD-065500) note at 47:

> The existing empirical evidence concerning the effectiveness of traditional cognitive and affective education approaches to tobacco, alcohol, and drug abuse prevention have consistently indicated that these approaches are *not* effective. There is even some evidence that approaches which attempt to dissuade adolescents from becoming involved with drugs by providing them with factual information concerning the consequences of drug use may actually lead to increased usage, possibly because they may serve to stimulate curiosity. (Citations omitted.)

c. THE SOCIAL INFLUENCES APPROACH.

Q. What was the next major approach to adolescent smoking prevention identified by the 1994 Surgeon General's Report?

A. The "social influences" approach, which attempted to teach adolescents how to resist perceived "social influences" toward smoking, such as peer pressure or cigarette advertising. *See, e.g., School Programs To Prevent Smoking,* National Cancer Institute Guide to Strategies that Succeed (1990) (JD-060921) at Chapter 2, "Strategies" (describing the social influences approach as: "Include information about social influences on tobacco use and about tobacco's short-term effects on the body. Teach students *how* to refuse.") (emphasis in original).

Interventions implementing the affective education model and the social influences model often have similar components. The line of demarcation between them is not always crystal clear and distinct to me.

Q. Did interventions implementing the social influences approach to youth smoking prevention have, as a component of the intervention, an effort to impart additional information on the health risks of smoking?

A. Many of them did. For example, although the CDC's *Guidelines for School Health Programs To Prevent Tobacco Use and Addiction* (U.S. Department of Health and Human Services, Public Health Service, (1994) (JD-001630) recognized (at 8) that "providing only factual information about the harmful effects of tobacco use" does "not prevent tobacco use, may stimulate curiosity about tobacco use, and may prompt some students to believe that the health hazards of tobacco use are exaggerated," it nonetheless included as part of the Guidelines, providing information about the "[i]mmediate and long-term undesirable physiologic, cosmetic, and social consequences of tobacco use."

Q. Did interventions implementing the social influences approach, which often had an information-based component, successfully prevent youth smoking initiation?

A. They had essentially no effect.

The most rigorous scientific evaluation of the efficacy of the social influences approach that I found was Peterson et al., "Hutchinson Smoking Prevention Project: Long-Term Randomized Trial in School-Based Tobacco Use Prevention -- Results on Smoking," *J. Nat'l Cancer Inst.*, 92(24):1979-1991 (2000) (JD-062758).

The Hutchinson Smoking Prevention Project (HSPP) was an intensive school-based intervention. It met the guidelines for planning and implementing school-based programs to

prevent tobacco use that were recommended by the Centers for Disease Control and Prevention, including as part of the intervention an attempt to build motivations based on the "general" and "long-term" health effects of smoking (JD-062758 at Table 1, page 1982). It lasted for 15 years. It involved 40 school districts that were first matched and then randomized to receive or not receive the intervention. The intervention itself started in third grade and continued through twelfth grade. The dropout rate was less than six percent. The correct unit of analysis was employed. In short, it was methodologically superb; in the words of the editorial accompanying its publication, it represents the "gold standard" methodologically.

Peterson et al. (JD-062758 at 1979) showed no effects at all of the intervention:

> The rigor of the HSPP trial suggests high credence for the intervention impact results. Consistent with previous trials, there is no evidence from this trial that a school-based social-influences approach is effective in the long-term deterrence of smoking among youth.

Some less well-conducted studies involving social influences interventions have suggested that the interventions may delay the onset of smoking for a short time, but generally they have found that the interventions failed to affect smoking behavior over a longer period of time, such as at the time of high school graduation.

As the 2000 Surgeon General's Report (U.S. Exhibit 64,316) noted (at 71 and 73):

> [P]rogram effects on cigarette smoking and other substance use behaviors had disappeared at this time (one year after the end of the program). . . .
> By the end of high school, the program had no detectable effect on cigarette smoking. . . . Similar to the other longer-term followup studies, these outcomes indicated that program effects eroded rapidly when the program ended and that no effects on smoking behavior or related beliefs were detectable at a later time.
>
> * * * *
>
> The primary limitation of this promising record of success is its generally short-lived nature. Three of the studies that

> followed participants through the 12th grade consistently found that effects had faded over the high school years.

Murray et al., "The Prevention of Cigarette Smoking in Children: A Comparison of Four Strategies," *J. Applied Social Psychol.*, 14(3):274-288 (1984) (JD-064027) at 274 (abstract):

> These data suggest that a program which teaches specific skills to resist social pressures to begin smoking and which teaches students about the short term physiological consequences of smoking is more effective than a program which concentrates on long term health consequences. Perhaps most important, the use of same-age peer leaders as teachers appears to be a necessary condition for successful use of this intervention program.

Murray et al., "The Prevention of Cigarette Smoking in Children: Two- and Three-Year Follow-up Comparisons of Four Prevention Strategies," *J. Behav. Med.*, 10(6):595-611 (1987) (JD-064028) at 595 (abstract):

> Recent studies have suggested that a prevention program that addresses the social influences that encourage smoking can be effective in deterring cigarette use by adolescents. This study presents 1-, 2-, and 3-year follow-up results from two studies which evaluated three variations of the social influences curriculum and compared them to a health consequences program and a usual-care comparison group. These results suggest that a peer-led, social influences program can restrain smoking among both baseline nonsmokers and baseline experimental smokers at 2 years postintervention.

Murray et al., "Results from a Statewide Approach to Adolescent Tobacco Use Prevention," *Preventive Med.*, 21:449-472 (1992) (JD-063864) at 449 (abstract):

> *Results*. The prospective study indicated that none of the interventions was more effective in reducing adolescent tobacco use compared with a randomized control group. The serial cross-sectional study revealed that there was a modest net decline in Minnesota relative to Wisconsin from 1986 to 1990, but that it was within the range of chance variation.
>
> *Conclusions*. Taken together, these results indicate that this legislative initiative was insufficient to reduce adolescent tobacco use statewide during the 5-year study period. Together with results from other recent studies, they suggest that even more intensive efforts may be required to effect widespread reductions in adolescent tobacco use.

Flay et al., "Six-Year Follow-up of the First Waterloo School Smoking Prevention Trial," *Am. J. Pub. Health*, 79(10):1371-1376 (1989) (JD-064033) at 1371 (abstract):

> Abstract: This paper reports six-year follow-up data from the first large-scale randomized trial of the social influences approach to smoking prevention. In 1979, 22 schools were randomly assigned to program or control conditions. Students in program schools received a social influences curriculum in six core and two maintenance sessions in grade 6, two booster sessions in grade 7, and one booster session in grade 8. All students were assessed at pretest (T1), immediate posttest (T2), end of grade 6 (T3), beginning and end of grade 7 (T4 and T5), end of grade 8 (T6), and grades 11 and 12 (T7 and T8). Ninety percent of study students were relocated and data obtained from over 80 percent of them at T8. Program effects on experimental smoking observed in grades 7 and 8 had completely decayed by T8, six years after the beginning of the program. Grade 6 smoking experience and social risk were each strong predictors of T8 smoking behavior. Subjects who had left school were smoking at more than twice the rate of subjects still in high school (grade 12) at T8. We discuss implications of the results.

Q. Did any of the social influences interventions report that they successfully reduced smoking prevalence over the long-term, such as at high school graduation?

A. Some studies report that the Life Skills Training program, which teaches adolescents job search skills, listening skills, communication skills, and coping skills, such as anger management, can have an effect on smoking initiation through high school. However, as Dr. Biglan testified at trial (*U.S.* Trial Tr. 1/11/05 a.m. session, 9645:22-9646:3 and 9648:3-22), the results from the Life Skills Training program appear to be sensitive to differences in implementation, which is consistent with Tobler's observation (JD-063898 at 315 and Figure 5 at 311, discussed earlier) that, as the size of an intervention approaches a few thousand students, the efficacy of the interventions diminishes substantially.

d. MASS MEDIA CAMPAIGNS.

Q. Apart from school-based interventions, did you also encounter literature that attempted to assess media-based efforts to prevent smoking initiation by providing additional health-risk information?

A. Yes, although I did not attempt to be comprehensive, I did review some studies. And, as with the school-based interventions, I attempted to focus my attention on those media campaigns that conveyed additional information about the diseases and conditions caused by smoking. Media campaigns that communicated a different message, such as vilifying the cigarette manufacturers, and that did not have as a major element providing additional factual information about the health risks of smoking were not pertinent to this part of my work.

Also, as with the school-based interventions, I assessed the statistical reliability and validity of the reported results and weighted more heavily those studies that were methodologically superior.

Q. What did you find?

A. The better studies -- those that were randomized with long term follow-up, that had low dropout rates, and that were analyzed in a statistically appropriate way -- consistently support the conclusion that using the media to present additional health-risk information, beyond that which is otherwise readily available, has essentially no effect on youth smoking behavior. As Flay, "Youth Tobacco Use: Risks, Patterns, and Control," Chapter 19, *Nicotine Addiction: Principles and Management* (Orleans & Slade Eds.) New York: Oxford University Press (1993) (JD-065592) summarized (at 372):

> Tobacco education traditionally has consisted of information about the negative health consequences of tobacco use. While such an approach is effective in altering knowledge it typically has not effectively prevented onset. This is true both for programs imparted to school children in classrooms (Goodstadt 1978;

> Thompson 1978; Tobler 1986) and for broader campaigns directed to the general public through mass media (Flay 1987). Information is rarely, if ever, sufficient to alter behavior (Flay et al. 1983), regardless of how it is delivered.

Among others, Bauman et al., "The Influence of Three Mass Media Campaigns on Variables Related to Adolescent Cigarette Smoking: Results of a Field Experiment," *Am. J. Pub. Health,* 81(5):597-604 (1991) (JD-064065) reported (at 597) no detectable effect on smoking behavior; Murray et al., "Effects of a Statewide Antismoking Campaign on Mass Media Messages and Smoking Beliefs," *Prev. Med.*, 23:54-60 (1994) (JD-065744) found (at 54) that a dramatic increase in exposure to anti-smoking messages via the media had little effect on smoking behavior; Sowden et al., "Mass media interventions for preventing smoking in young people," *The Cochrane Library*, Issue 2, Chichester, UK: John Wiley & Sons, Ltd. (2004) (JD-065871) reports (at 1) that only 2 out of 6 controlled trials of media-based programs showed an effect and concludes "There is some evidence that the mass media can be effective in preventing the uptake of smoking in young people, but overall the evidence is not strong"; Thompson, "Smoking Education Programs 1960-1976," *Am. J. Pub. Health*, 68(3):250-257 (1978) (JD-060939) found (at 250) "little success" with information-based interventions presented through the media and by other means.

Finally, a community-based, randomized smoking cessation program, the COMMIT study, is also relevant to this question. Hymowitz, N. et al., "Predictors Of Smoking Cessation In A Cohort Of Adult Smokers Followed For Five Years," *Tobacco Control*, 6(suppl 2):S57-S62 (1997) (JD-001162); and The COMMIT Research Group, "Community Intervention Trial for Smoking Cessation (COMMIT): I. Cohort Results from a Four-Year Community Intervention," *Am. J. Public Health*, 85(2):183-192 (1995) (JD-062352).

COMMIT included a media campaign that communicated incremental, additional health-risk information in an effort, primarily, to increase smoking cessation among adult, heavy smokers (JD-062352 at 191, Appendix A). An examination of the results indicates that the media campaign had no effect on youth smoking prevalence as recently reported in Bowen et al., "Intervention effects on youth tobacco use in the community intervention trial (COMMIT)," *J. Epidemiol. Community Health*, 57:159-160 (2003) (JD-065557) at 160:

> The data reported here do not support the hypothesis that the adult focused COMMIT intervention was efficacious in reducing the prevalence of regular youth smoking. Among ninth graders living in treatment communities as well as among their counterparts living in comparison communities, the general trend was toward little or no difference over the time interval assessed (1990 to 1992) -- a leveling off in tobacco use rates that is consistent with national trends reported in other surveys conducted during this time period.
>
> * * * *
>
> It appears that the COMMIT intervention, which did target adult smokers, was not a cause of change in adolescent smoking behaviour.

Q. Is your opinion about the ineffectiveness of media campaigns, which provided additional health-risk information to curtail youth smoking, shared by other reviewers?

A. Yes. The 2000 Surgeon General's Report (U.S. Exhibit 64,316) stated (at 79):

> Results indicated that the campaigns had effects on the recipients' knowledge of the consequences of smoking and other mediators but not on cigarette smoking behavior. (Citation omitted.)

Also, Pechmann, C., "Changing Adolescent Smoking Prevalence: Impact of Advertising Interventions," Chapter 10, *NCI Monograph 14: Changing Adolescent Smoking Prevalence* (2001) (U.S. Exhibit 72,977), reports (at 171):

> To date, there is little conclusive evidence of a direct link between advertising-only interventions and reduced adolescent smoking prevalences.

e. CONCLUSION: PROVIDING ADDITIONAL HEALTH RISK INFORMATION, ABOVE AND BEYOND THE INFORMATION ALREADY AVAILABLE, IS "NOT EFFECTIVE" IN PREVENTING SMOKING INITIATION.

Q. Professor Rubin, based on your review and assessment of the literature, do you have an opinion on whether providing adolescents with additional information about the long-term health risks of smoking is effective in preventing smoking initiation?

A. I do.

Based on my systematic review, it is clear that merely providing additional information about the diseases and conditions caused by smoking is not effective in reducing the prevalence of smoking through high school graduation.

Q. Is there evidence of a long-term effect of information-based interventions on adolescent smoking initiation in any subpopulation?

A. None that I saw.

Q. Are you alone in holding the opinion that providing additional information about the health risks of smoking is ineffective in preventing long-term smoking initiation?

A. No. Based on my review, nobody really disagrees. I have compiled in a summary exhibit, JD-068090, a set of quotations from the literature in which various authors express views consistent with my own opinion. Here is one example:

Swisher et al., "Information: The Irrelevant Variable in Drug Education," Chapter 4 in *Drug Abuse Prevention Perspectives and Approaches for Educators*, Corder, Smith and Swisher, Wm. C. Brown Co. Publishers (1975) (JD-068085) report (at 58 and 60, respectively):

> Again, this finding lends support to the notion that drug information in itself is insufficient to affect students' attitudes and use of drugs.
>
> * * * *

> Drug education programs of the information type may not be as effective as we have hoped that they would be. Indeed, we may have discovered that information is the irrelevant variable.

Moreover, the Government's own experts have given testimony that is consistent with my opinion. Dr. Biglan, for example, agreed with the Surgeon General that merely providing additional information about the long-term health effects of smoking was ineffective in reducing adolescent smoking initiation. (*U.S.* Trial Tr. 1/11/05 A.M. session , 9642:11-9644:19 and 9657:13-9658:15).

Q. Are you saying that health-risk information cannot affect youth smoking initiation?

A. No. I do not disagree, for example, with the view expressed in the 1989 Surgeon General's Report that the 1964 Surgeon General's Report, which concluded that cigarette smoking caused lung cancer in men, and the subsequent anti-smoking campaign led to decreased youth smoking prevalence.

Rather, the studies that I have examined have investigated a different question: whether *incremental*, additional information about the health risks of smoking, over and above that which was otherwise readily available to essentially everybody, had a causal effect on youth smoking information. The answer seems very clear: such information had essentially no effect.

2. CAN THE ABSENCE OF A CAUSAL EFFECT ON ADOLESCENT SMOKING INITIATION RESULTING FROM PROVIDING ADDITIONAL, INCREMENTAL TRUTHFUL INFORMATION ON THE HEALTH RISKS OF SMOKING BE USED TO INFER THE CAUSAL EFFECT OF DEFENDANTS' ALLEGED SUPPRESSION OR MISREPRESENTATIONS OF HEALTH RISKS OF SMOKING ON ADOLESCENT SMOKING INITIATION?

Q. Based on the studies showing that extra information about the health risks of smoking essentially show no effect on smoking initiation, what, if anything, have you concluded about the effect of the Defendants' alleged information-based misconduct on smoking initiation?

A. I conclude that the alleged information-based misconduct of the Defendants -- the alleged suppression of health risk information and their past statements that smoking had not been proven scientifically to cause diseases, such as lung cancer -- had essentially no effect on youth smoking initiation.

Q. Have there been any studies that examined specifically the effect on smoking initiation of either Defendants' alleged suppression of health-risk information or of Defendants' past statements that smoking had not been proven scientifically to be a cause of lung cancer and other diseases?

A. No.

Q. How is it possible to draw inferences about the effect on youth smoking initiation of the Defendants' alleged information-based misconduct concerning the health risks of smoking if it has never been assessed directly?

A. Generally, as we discussed earlier, one of the central tasks in the field of causal inference is to estimate what would have happened in a counterfactual world that never existed. It is precisely what statisticians and others interested in causal effects do professionally.

To do so, one always has to rely upon data from the actual world as it existed, together with assumptions about how actual world data inform us about the counterfactual world. It is not unusual to have a situation, such as the one here, where the actual world data do not directly estimate the effect of the exact same intervention as the one about which we want to draw a causal inference. Recall that I also gave some examples of "bridging" from my own work with the CDC on anthrax vaccine from non-human primates to humans, the "bridging" from adult data to children for Novartis' epilepsy medication, and the bridging in time based on the observational study comparing mastectomy and breast conservation surgery.

Here, the evidence is very strong that providing additional, incremental truthful information about the health risks of smoking makes essentially no difference to youth smoking initiation. Providing adolescents with additional information about the long-term health risks of smoking, above and beyond that which was already available, has been shown to be ineffective in reducing smoking initiation. Studies assessing this were conducted as early as the 1960s and continued for decades, against an increasing background level of health-risk information concerning smoking.

For assessing the causal effect of the defendants' alleged information-based misconduct, one must "bridge" from (a) this lack of causal effects from incremental additional truthful health information to estimate (b) causal effects of the reduced truthful health information due to the defendants' alleged misconduct.

Q. What is the result of your assessment following that approach?

A. The evidence since the 1950s strongly supports the inference that the defendants' alleged information-based RICO violations concerning the long-term health risks of smoking had essentially no effect on adolescent smoking initiation in any subpopulation.

B. DID FEWER ADULTS QUIT SMOKING BECAUSE OF DEFENDANTS ALLEGED FRAUD CONCERNING DISEASES AND CONDITIONS CAUSED BY SMOKING?

Q. Let's turn to smoking cessation among adults. Were you asked to assess the causal effect of Defendants' alleged information-based RICO violations -- their alleged suppression and misrepresentations of the health risks of smoking -- on smoking cessation?

A. Yes.

Q. Please describe generally what you did.

A. Just as I did in my assessment of the effect of this alleged misconduct on adolescent smoking initiation, I reviewed the literature on the effect of providing incremental, additional

truthful information about the health risks of smoking, over and above that which was already available, on smoking cessation, focusing on those studies that were more statistically reliable and valid.

Q. What did you find?

A. There are very few interventions for smoking cessation that were purely information-based, but a number of interventions had an information-based component. Perhaps this is because principally information-based programs, such as self-help approaches, which typically included or focused on information about the health risks of smoking, are generally understood to have essentially no effect on smoking cessation.

For example, of the hundreds of interventions reviewed by Fiore et al. (1996, 2000) (JD-063828 and JD-001210) and including those contained in the Cochrane Database of Systematic Reviews at U.S. Exhibit 89,465; U.S. Exhibit 89,466; U.S. Exhibit 89,467; U.S. Exhibit 89,468; and U.S. Exhibit 89,469 which were designed by health-care professionals to increase smoking cessation, why didn't many of them simply provide additional information about the health-risks of smoking if that were all that was needed to increase quitting? It seems clear that, as with youth initiation, "information is the irrelevant variable."

Nevertheless, some studies of self-help materials, which can be regarded as having a substantial information-based component, do exist and they show no real effect of the intervention. For instance, Lando et al., "A Comparison of Self-Help Approaches to Smoking Cessation," *Addictive Behav.*, 16(5):183-193 (1991) (JD-068084) report (at 183 (abstract) and 190, respectively):

> Results at 7-month follow-up failed to indicate treatment effects either for abstinence or for reported quit attempts.

* * * * *

> Results failed to indicate any difference in effectiveness between the Quit and Win and Quit for Good self-help materials. Contrary to expectations, neither set of self-help materials led to higher abstinence rates than were achieved by control subjects who did not receive materials.

In 1993, McFall et al., "The Effects and Use of Maintenance Newsletters in a Smoking Cessation Intervention," *Addictive Behav.*, 18(2):151-158 (1993) (JD-065718) observed (at 151 (abstract) and 156, respectively):

> The maintenance condition did not increase cessation at any wave of interviewing, assessed by multiple point or point prevalence of abstinence.
>
> * * * *
>
> The maintenance materials were not associated with higher rates of cessation or lower rates of relapse among those abstinent at the end of the program.

Finally, as I noted earlier, the COMMIT randomized trials used a community-level four-year intervention designed to increase quit rates of heavy smokers (JD-001162 and JD-062352). The intervention was multifaceted. Among other things, it consisted of mandated:

> Mass media campaigns promoting smoking as a public health problem and promoting cessation.
>
> Sponsored "magnet events" such as "quit and win" contests and the Great American Smokeout.
>
> Comprehensive training in smoking cessation techniques for health-care providers, such as physicians and dentists, to share with their patients.
>
> Work site annual smoking workshops, promotion of the magnet events, incentives for smoking cessation, and distribution of smoking cessation materials.
>
> Newsletters and information mailed to smokers about smoking cessation resources.

(JD-062352 at 191, Appendix A.)

COMMIT had no effect on smoking cessation among the target population, heavy smokers. Although there was a small increase in quitting (less than 1% per year) among light to

moderate smokers, they found essentially no effect on overall smoking prevalences, presumably because of initiation and relapse.

Q. Is your opinion consistent with views expressed in the various review articles that you examined?

A. Yes, including two relatively recent reviews, one which was authored by one of the Government's experts and the other which was published by the Cochrane Database of Systematic reviews.

In 2000, Fiore et al., *Clinical Practice Guideline: Treating Tobacco Use and Dependence*, U.S. Department of Health and Human Services (2000) (JD-001210) report (at 63) that "[t]he previous analyses failed to show a consistent, beneficial effect due to self-help" and summarized (at 64, Table 19) a meta-analysis of 21 self-help studies that showed no statistically significant effects on abstinence rates.

In 2002, Lancaster & Stead, after surveying 23 different randomized trials, conclude (at 1-2) that, although standard self-help materials "may" increase quitting, any effect "is likely to be small." (Lancaster & Stead, "Self-Help Interventions for Smoking Cessation," *The Cochrane Database of Systematic Reviews*, Issue 3. Art. No.: CD001118 (2002) (U.S. Exhibit 89,466).)

These more recent reviews, moreover, are consistent with earlier review articles.

In 1997, Klausner, R., "Evolution of Tobacco Control Studies and the National Cancer Institute," *Tobacco Control* 6(Suppl 2): S1-S2 (1997) (JD-061644) wrote (at page S-1):

> Findings from over 100 intervention trials continue to provide much new and important information about how to reach smokers and potential smokers. However, one of the major conclusions that one can draw from these studies is that large-scale reductions in smoking prevalence are unlikely using interventions that are primarily directed toward the individual and delivered through traditional intervention channels.

In that same year, Ebrahim and Smith, "Systematic Review of Randomised Controlled Trials of Multiple Risk Factor Interventions for Preventing Coronary Heart Disease," *British Med. J.* 314:1666-1674 (1997) (JD-063824) wrote (at 1666):

> More recent trials examining changes in risk factors have cast considerable doubt on the effectiveness of these multiple risk factor interventions and even interventions against smoking, prompting a review of the reasons for the frequent failure of such community experiments.

Finally, Fiore et al., *Smoking Cessation: Clinical Practice Guideline No. 18*, USDHHS, Public Health Service, Agency for Health Care Policy and Research, AHCPR Pub. No. 96-0692 (1996) (JD-063828), summarized 300 randomized trials, suggesting modest effects, at best, of many interventions, and essentially no reported effect for the type of intervention that could serve to "bridge" to an estimate of the causal effect of the tobacco industry's alleged, information-based RICO violations on smoking cessation. It stated (at 1-2) in the executive summary that three treatment elements are effective in promoting smoking cessation, and providing incremental, additional health-risk information is not among them:

> Three treatment elements, in particular, are effective, and one or more of these elements should be included in smoking cessation treatment:
>
> - Nicotine replacement therapy (nicotine patches or gum)
> - Social support (clinician-provided encouragement and assistance)
> - Skills training/problem solving (techniques on achieving and maintaining abstinence).

Q. Do you have an opinion about the effect of the Defendants' alleged information-based misconduct -- their alleged suppression and misrepresentations about the health risks of smoking -- on smoking cessation?

A. Yes. I "bridge" from (a) this lack of evidence for any causal effects on smoking behavior from incremental, additional truthful health information on smoking cessation in the recent past decades to estimate (b) the causal effects on smoking cessation of the reduced truthful health information due to the defendants' alleged misconduct in the more distant past. I conclude that the Defendants' alleged information-based misconduct concerning the health-risks of smoking had essentially no effect on smoking cessation.

C. ARE THE CAUSAL ASSERTIONS OF DR. BURNS CONCERNING THE EFFECT OF DEFENDANTS' ALLEGED MISREPRESENTATIONS OF THE HEALTH RISKS CAUSED BY SMOKING BASED ON STATISTICALLY VALID AND RELIABLE EVIDENCE?

Q. Dr. Burns, in his written direct examination (at 71:13-19), asserted that "it is clear that the misconduct of the tobacco industry in denying the disease risks caused by smoking . . . has led to substantially higher rates of cigarette consumption than would have occurred absent these actions." Is this assertion based on statistically valid and reliable evidence?

A. In the course of my review, I found no statistically valid or reliable evidence that could be used to support this opinion. And, as Dr. Burns testified on cross examination (Burns, *U.S.* Trial Tr. 2/16/05, at 13591:22-13592:13), he presented no such evidence to support his opinion in either his written direct examination or his expert report:

> Q. Isn't it true that you present no quantitative evidence or estimate of the effect of the defendants' assertions that smoking had not been proven to cause disease on smoking prevalence?
>
> A. That's correct.
>
> Q. And you have not made and produced such a quantitative estimate during the course of your work in this case, correct?
>
> A. That's also correct.

Q. In fact, you cite nothing in your written direct examination at pages 69 to 71 to support even your non quantitative assertions about the effects of defendants' statements on whether or not smoking caused lung cancer on the prevalence of smoking, right?

A. That is correct.

Q. You don't cite to any peer reviewed literature on the subject, correct?

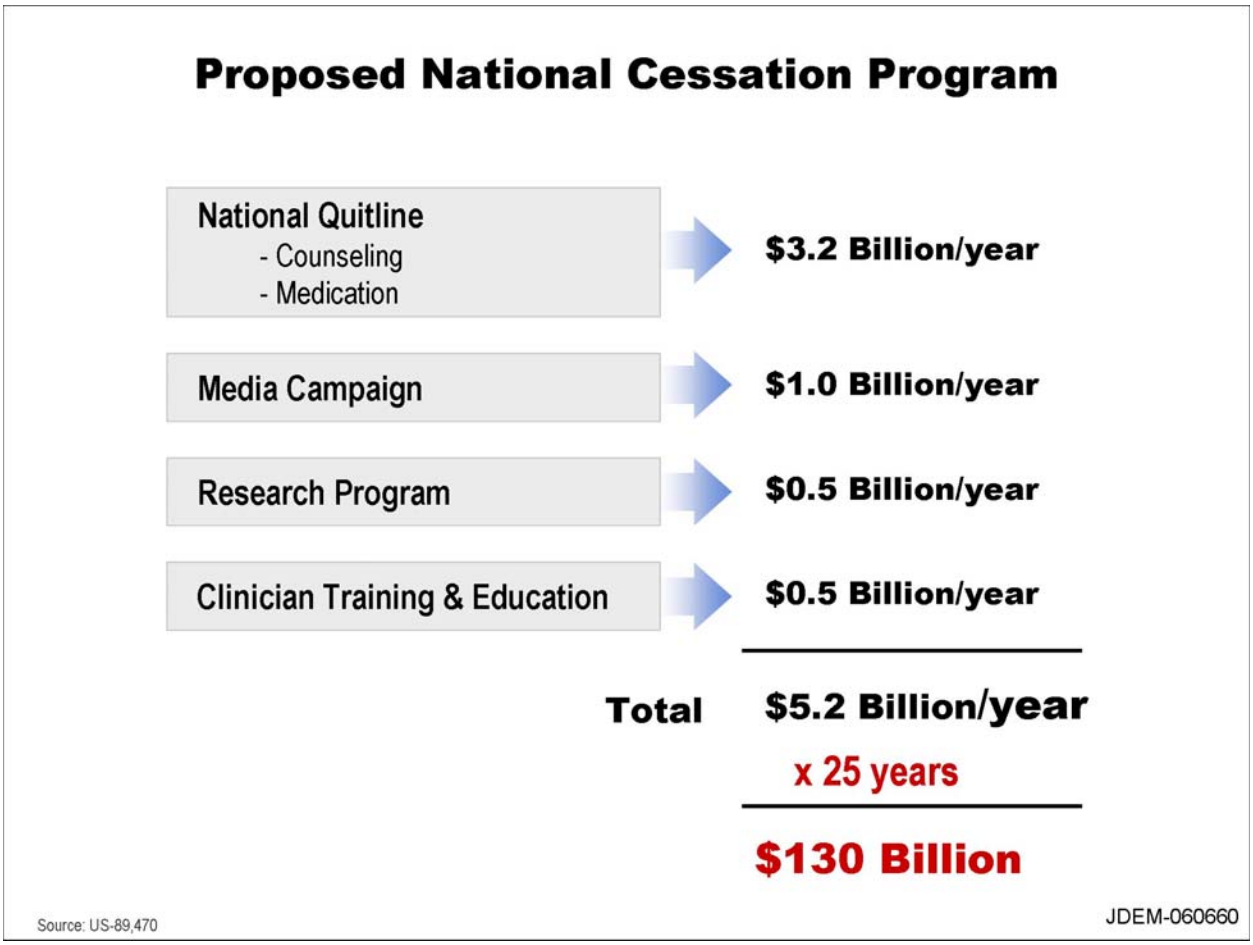
A. Don't cite anything.

V. THERE IS A STATISTICALLY RELIABLE AND VALID MANNER TO ESTIMATE THE CAUSAL EFFECTS FROM A NATIONAL SMOKING CESSATION PROGRAM SUCH AS THE ONE DESCRIBED BY DR. FIORE, BUT DR. FIORE HAS NOT EVEN ATTEMPTED TO ESTIMATE THOSE CAUSAL EFFECTS RELIABLY OR VALIDLY.

A. THE COST AND EFFECTIVENESS OF THE PROPOSED NATIONAL CESSATION PROGRAM.

Q. What are the parts of Dr. Fiore's proposed national cessation program?

A. As outlined Dr. Fiore's expert report (at 4-7) and as illustrated in JDEM-060660, the program has four major parts: (a) a national quitline, (b) a media campaign, (c) a research program, and (d) a clinician training and education program.



Q. I'd like to first look at the cost of the national cessation program and then at its impact on smoking cessation.

A. Fine.

B. THE COST OF THE NATIONAL CESSATION PROGRAM.

Q. As proposed, what would each part of the program cost?

A. Their respective costs are, as reflected in the prior demonstrative, proposed to be $3.2 billion per year for the national quitline, at least $1 billion per year for the media campaign, $500 million per year for the research program, and $500 million per year for the clinician training and education program. That's $5.2 billion per year.

Q. How much would the program cost over its life?

A. Dr. Fiore's recommends that the program continue at that funding level for approximately 25 years. (*See* Fiore Written Direct at 69:1-19; Fiore Expert Report at 9; Fiore *U.S.* Dep. 5/5/05 at 54:5-14.) Accordingly, the total cost of the proposed national cessation program would be about $130 billion.

Q. Do you have an opinion whether it is possible to estimate the costs of the proposed national cessation program in a statistically valid and reliable manner?

A. Yes, I do.

Q. What is your opinion?

A. It is possible to estimate the costs of the proposed national cessation program in a statistically valid and reliable way, using actual world data, particularly data from already-established state quitlines, coupled with explicit assumptions.

Q. Based on your review of Dr. Fiore's expert report, deposition testimony, and written direct testimony, as well as certain underlying materials he considered, do you have an opinion on whether the cost estimates of any of the components of Dr. Fiore's proposed national cessation program are statistically valid and reliable?

A. Yes, I do.

Q. What is your opinion?

A. Based on what I have seen, the costs of the components of the national cessation program proposed by Dr. Fiore are not based on statistically valid or reliable analyses of data.

Q. Let's look at each of the program's parts.

A. Okay.

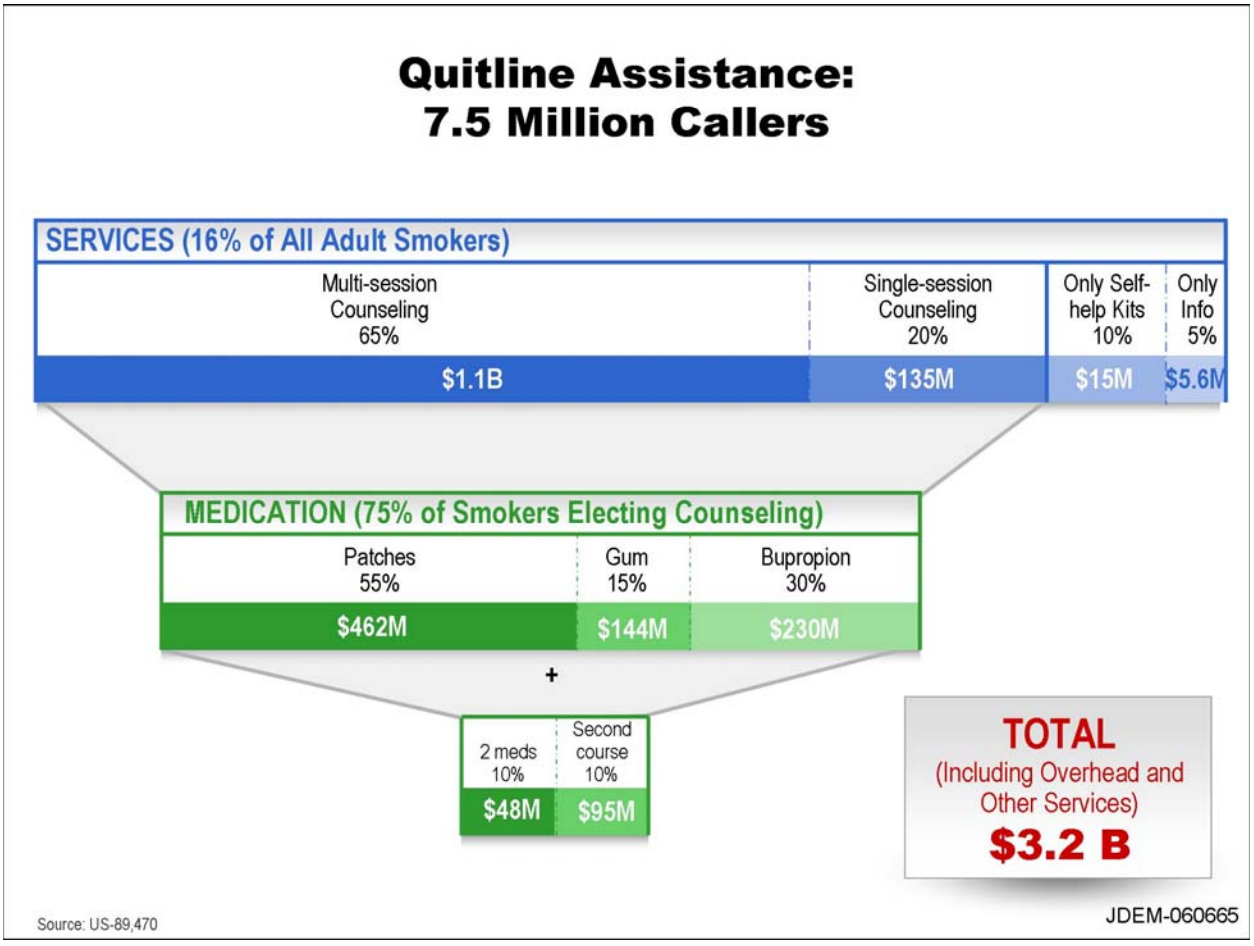
1. NATIONAL QUITLINE.

Q. I'd like to start with the national quitline. What are the major factors that need to be estimated in order to estimate the cost of the proposed national quitline?

A. A major factor driving the cost of the proposed national quitline is the extent to which smokers will actually use it -- the participation rate.

Q. A demonstrative, JDEM-060665, summarizes how the costs of the proposed national quitline flow from the participation rate. Is this consistent with your understanding based on reading Dr. Fiore's written direct examination and deposition testimony?

A. Yes, JDEM-060665 is consistent with my understanding based on U.S. Exhibit 89,470 (at VXA 452 0051).

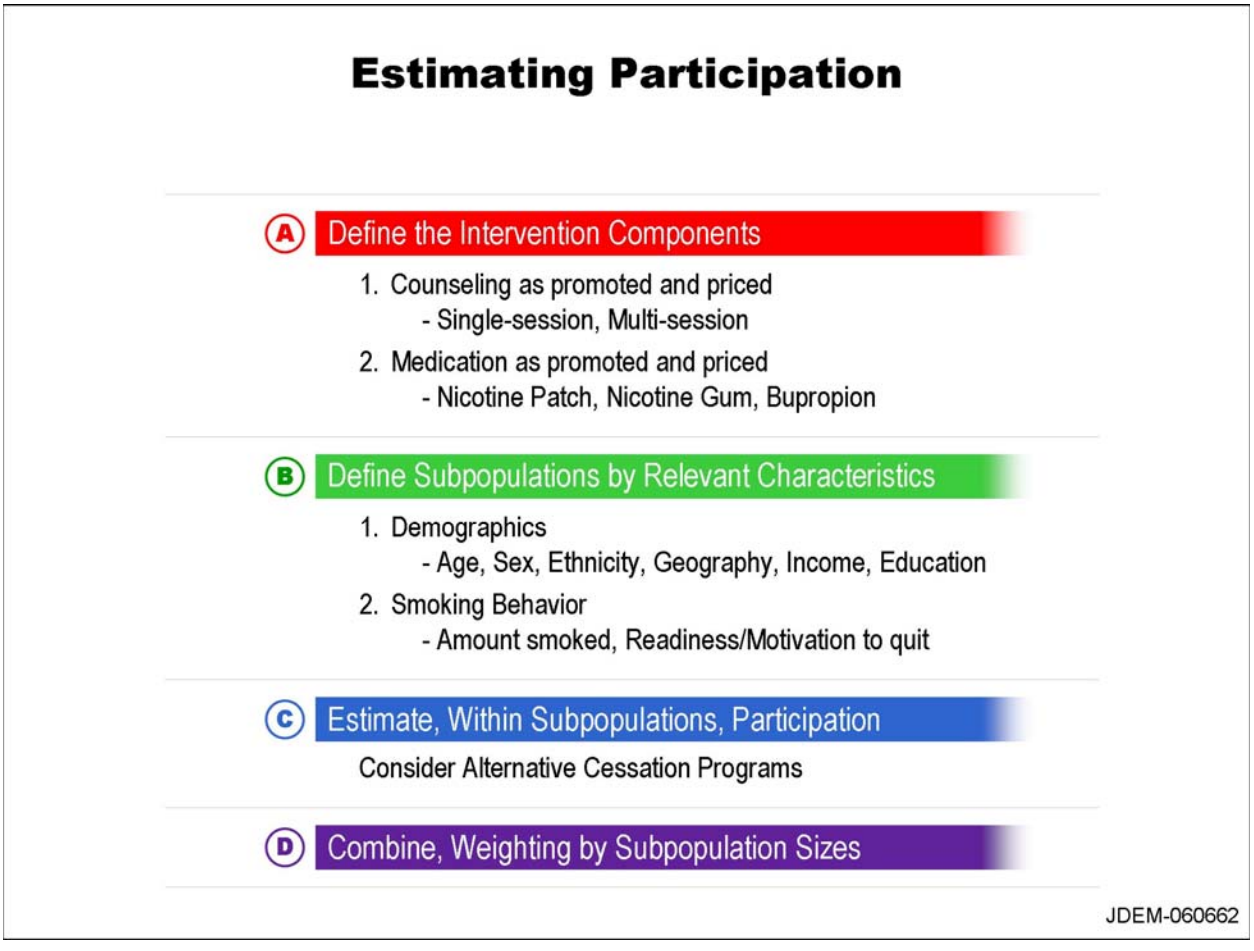


As summarized in the demonstrative being displayed, Dr. Fiore starts from the premise that 16% of U.S. adult smokers will call the proposed national quitline. Moreover, among those calling, he assumes they will use the different cessation interventions offered through the

proposed national quitline at a variety of different rates that, when multiplied by estimated unit costs, generate a cost estimate of over $2.3 billion per year for cessation counseling and medication. (*See* U.S. Exhibit 89,470 at VXA 452 0051.) I also understand that, after adding in other services, research, and overhead, the annual total cost that Dr. Fiore states for the proposed national quitline is about $3.2 billion. (*Id.* at VXA 452 0051-0052.)

Q. Before we address how Dr. Fiore arrived at a 16% participation rate for the proposed national quitline, how should one go about estimating the participation rate for a national quitline such as described by Dr. Fiore?

A. The general approach is depicted in JDEM-060662.



Initially, we would need to identify the various intervention components of the national quitline. The major intervention components in Dr. Fiore's national quitline are (a) self-help materials; (b) telephone-based counseling, either single-session counseling or multi-session counseling; and (c) medications, including nicotine patches, nicotine gum, bupropion, and sometimes multiple medications or multiple courses of medication. In addition to identifying these various components, we would need to define (i) how they are to be offered (*e.g.*, their price and the level of promotion); and (ii) whether they are otherwise available through, for example, existing state quitlines, Medicaid, Medicare, or private insurance (*i.e.*, whether potential program users may use other programs and treatments).

Q. What would you need to do next?

A. We would need to partition the target population (*i.e.*, smokers), both demographically and by relevant smoking behavioral characteristics, into relevant subpopulations.

Q. Why is that important?

A. For the same kinds of reasons we discussed earlier; namely, different demographic groups appear to use smoking cessation aids at different rates. (*See* JD-068086 at 5-8 (discussing gender, age, race, and education of callers to Wisconsin's quitline in 2003).) Also, those subpopulations with different smoking behavior characteristics (*e.g.*, different levels of readiness or motivation to quit or different intensity levels of smoking) would presumably utilize smoking cessation aids at different rates. (*See* JD-068086 at 5-8.) For example, there is evidence that smoking cessation techniques or therapies that are utilized at one rate in a subpopulation of smokers who are ready and motivated to quit would be used at lower rates in subpopulations that are less ready and motivated to quit. (*See* JD-068086 at 8.) Similarly, nicotine replacement therapies that are utilized at one rate in a subpopulation of light or moderate smokers may be

used at a different, perhaps higher rate in a subpopulation of similarly-motivated, heavier smokers. Therefore, these subpopulations of the population are defined by these relevant demographic and smoking behavioral characteristics.

Q. Once you have identified the individual cessation components of the national quitline and partitioned the smokers into appropriate subpopulations, what would the next step be to estimate the number of participants in each component nationally?

A. We would, for each subpopulation, estimate the percent of smokers who would participate in the various components of the program. To obtain a final estimate of the participation rate in the components in the nation, we would then weight the estimated rates in each subpopulation by the subpopulation's relative size in the nation. Finally, to estimate how many smokers would call the quitline and choose to participate in the various components, we would multiply those national participation rates by the size of the smoking population in the United States.

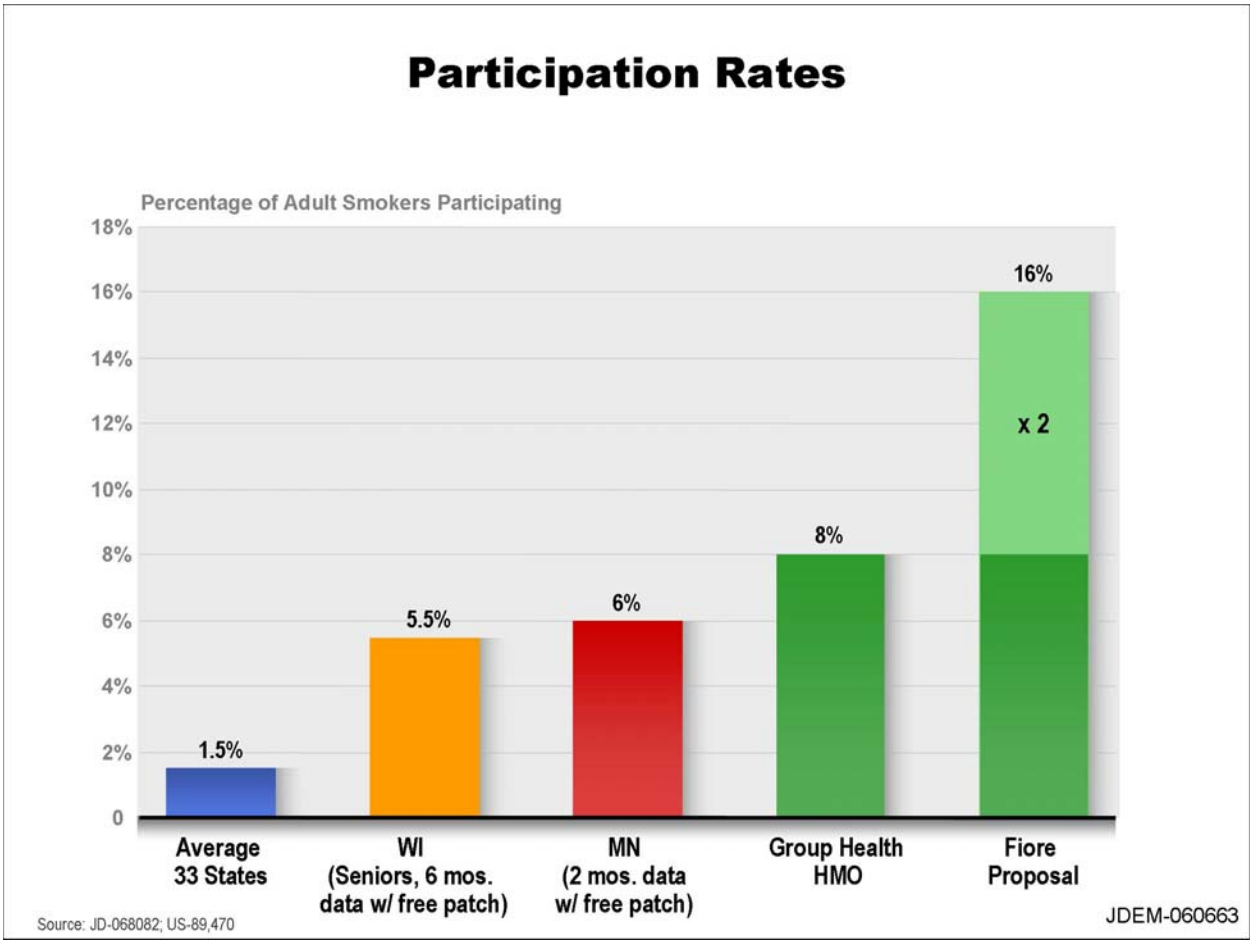
Q. Is the analytical process you just outlined the one used by Dr. Fiore to determine the anticipated participation rate for the proposed national cessation program?

A. Based on my review of Dr. Fiore's expert report, deposition testimony, and written direct trial testimony, his method did not resemble the analytical framework outlined above.

Q. What is your understanding of how Dr. Fiore arrived at a participation rate?

A. According to Dr. Fiore, the Cessation Subcommittee of the Interagency Committee on Smoking and Health that Dr. Fiore chaired received testimony in 2002 and early 2003 about participation rates. (Fiore *U.S.* Dep. 5/5/05 at 131:3–133:7; 134:6–139:6.) The Cessation Subcommittee received testimony that (a) according to a May 2002 survey, the 33 then-existing statewide quitlines had an average participation rate of 1.5% (JD-068082, at VXA 452 0431); (b)

based on six months of experience, Wisconsin had a 5.5% participation rate among senior citizens after free nicotine replacement therapies were made available to them (U.S. Exhibit 89,470, at VXA 452 0055); (c) based on two months of experience, Minnesota had an annualized 6% participation rate when it made free nicotine replacement therapies available through its quitline (*id.*); and (d) Group Health Cooperative, a HMO based in Seattle, Washington, had an 8% annual participation rate for the prior five years (*i.e.*, 1997 or 1998 through 2001 or 2002) in its HMO with no promotion, and "doubling that [participation] … rate to 16% does not seem unreasonable or unlikely" if a "national helpline would receive promotional and advertising funding at least equivalent to the funding allocated to provide helpline services" (*id.* at VXA 452 0058). This testimony is summarized in the bar graph that is JDEM-060663.



Q. What participation rate did the Cessation subcommittee decide to use?

A. According to Dr. Fiore, the Cessation Subcommittee, after some discussion, decided to adopt Group Health Cooperative's hoped-for 16% participation rate, which in turn is the national quitline participation rate that Dr. Fiore used in his testimony in this case. (*See* Fiore Written Direct at 52:7–53:14; *see also* Fiore *U.S.* Dep. 5/5/05 at 140:6-19; 143:6-10; 151:1-16.)

Q. In your opinion, has Dr. Fiore provided -- or did the Cessation Subcommittee provide -- a statistically valid or reliable basis for using a 16% participation rate?

A. No.

Q. Why do you say that?

A. Virtually none of the things that one must do to estimate the participation rate in a statistically valid or reliable manner appears to have been done. There is no indication that any effort was made to (a) define the various intervention components and their level of promotion and cost; (b) partition the smokers into relevant subpopulations; and (c) estimate an overall participation rate that properly weights the various underlying component and subpopulation-specific participation rates.

Instead, Dr. Fiore and the Cessation Subcommittee apparently discussed and then adopted Group Health's suggestion that the 8% participation rate from Group Health's HMO should be doubled to 16% because the quitline would be promoted. (*See* Fiore Written Direct at 52:7–53:14; *see also* U.S. Exhibit 89,470; Fiore *U.S.* Dep. 5/5/05 at 140:6-19; 143:6-10; 151:1-16.)

Q. Has any evidence been provided that the 8% participation rate Group Health said it achieved in its HMO is a statistically valid or reliable basis for estimating the participation rate for a national quitline?

A. To the best of my knowledge, no.

Q. Assuming that Group Health did achieve an 8% participation rate for its smoking cessation program in its HMO, are there reasons why their experience might not provide a statistically valid or reliable basis for extrapolating to the anticipated participation rate a national quitline such as the one Dr. Fiore proposes might achieve?

A. Yes, there are. First, and apart from the fact that the 8% is not broken down by either cessation components or subpopulations, it appears to cover an unrepresentative subpopulation of the United States. According to Group Health, its HMO insures roughly 600,000 members, largely in the Seattle, Washington area. (*See* U.S. Exhibit 89,470 at VXA 452 0055; JD-068080 at 1.) The participants in the Group Health HMO are insured and, in that respect, are not representative of the United States. (*See* Fiore *U.S.* Dep. 5/5/05 at 120:21-25 ("Q. . . . [A]s a general rule insured populations will have differences from the general demographics of the United States, correct? A. Correct.").)

Another nonrepresentative feature is that Group Health reports that the smoking prevalence in its HMO was about 15%, yet surveys show that the overall smoking prevalence in the United States was over 20%. (*Compare* JD-068080 at 12 (reporting smoking prevalence in Group Health's HMO at 14-15% from 1998-2000); JD-068083 at 6 (presentation by Group Health's Dr. Tim McAfee stating that, within Group Health system, adult smoking prevalence was 15%); *and* Fiore *U.S.* Dep. 5/5/05 at 128:18-21 (discussing JD-068083); *with* JD-012906 at 427 (reporting that U.S. smoking prevalence ranged from 24.1% in 1998 to 22.5% in 2002).)

And, according to U.S. census data, the county in Washington that includes Seattle had a higher per capita income and a higher average level of education than the United States as a

whole. Both of those factors are associated with smoking behaviors. (*See* JD-012906 at 427 (smoking prevalence is higher among persons with less income and less education).)

Second, there is other experience, including experience with statewide quitlines, that may be more directly relevant when attempting to estimate a nationwide participation rate.

Q. Have you seen any statistically valid or reliable evidence that supports doubling the 8% participation rate?

A. No, I have seen no evidence that Dr. Fiore or the Cessation Subcommittee performed a valid or reliable statistical analysis underlying their statement that a 16% participation rate would be achieved. Although more intense promotion may increase participation, no justification for doubling the maximum observed participation rate was provided.

Q. Are you aware of any statistically valid or reliable analysis that Dr. Fiore or Group Health conducted to estimate how the 16% participation rate might change over time?

A. I have seen no such analysis.

Q. Would such an estimate be important?

A. Yes. A new program such as that proposed by Dr. Fiore could, initially, draw on pent up demand for the various smoking cessation components in the manner they would be provided (*e.g.*, a previously unsatisfied demand for free counseling or for free medication) that, over time, might dissipate. In addition, if the national cessation program were to increase the overall cessation rate, the percentage of smokers ready and willing to attempt to quit might, over time, become smaller. And, if the overall rate of cessation were to increase, the starting point in Dr. Fiore's and Group Health's cost calculations -- the number of adult smokers in the United States -- presumably would decline. Therefore, the number of participating smokers would --

and, assuming constant per-smoker costs, the proposed budget presumably should -- decline even if the same participation rate could be maintained.

Q. There are existing smoking cessation programs. Did you see any analysis or consideration of the extent to which smokers might use various intervention components from sources other than the proposed national quitline (*e.g.*, existing state quitlines; existing counseling or medication programs covered by Medicare, Medicaid, or private health insurance)?

A. Again, I saw no such analysis.

Q. Why might that be important?

A. Even if we knew the rate at which the subpopulations in the population would use a particular type of intervention component in the national quitline, when that component is already available from existing sources, some of the participating users would presumably use the intervention component from sources other than the national quitline. Accordingly, to estimate the participation rate the national quitline would achieve, which is the quantity of interest here, we would need to account for the fact that, among those who would choose to participate in some smoking cessation program, only some would use the national quitline.

Q. Are you setting an impossibly high statistical standard for determining the participation rate for the proposed national quitline?

A. No, I'm not for at least three reasons. First, I've done analogous things in other contexts. Second, although the approach I have outlined would involve substantial work, it seems appropriate in light of the $130 billion expenditure Dr. Fiore proposes. Finally, there appears to be significant actual world data from which to estimate participation rates.

Q. You mentioned that you had done analogous things in the past. What were you referring to?

A. For over three decades, I have consulted for various parts of the federal government in designing surveys to help make national estimates, including the Internal Revenue Service, the Department of Labor, the Bureau of Labor Statistics, the Census Bureau, the National Center for Health Statistics ("NCHS"), the Centers for Disease Control ("CDC"), the General Accounting Office, the Department of Transportation, the Department of Defense, the Department of Education, and others.

For a more specific example, for many years, I have consulted with the NCHS, which is part of the CDC, regarding missing data issues to ensure valid national and local estimates in the NHANES survey (National Health and Nutrition Examination Survey). And I have been involved in a series of projects for the United States Census Bureau, often in support of improving survey methods and data analyses useful for estimating the undercount in the decennial census.

Q. In your experience, do the other federal agencies and departments that you have worked with make national estimates in the way that Dr. Fiore or the Cessation Subcommittee he chaired apparently made them?

A. No. The federal agencies and departments with which I have worked attempted to make statistically valid and reliable estimates.

Q. You mentioned that additional data from real-world experience are available that could be used for estimating the participation rate. What were you referring to?

A. I have not done a comprehensive review or survey, but I have seen other participation rates. For example, Dr. Ossip-Klein et al. reported in an article published in 2003 that a

"nonrandom sampling of several of the established quitlines in North America" found that "1.1% to 1.7% of adult smokers called a quitline over the course of a year." (JD-068087 at 202.) And the CDC reports that, in 2001, 28 statewide quitlines had participation rates of between 1% and 5%. (JD-068088 at 18.)

Q. Were you referring to anything else?

A. As I mentioned before, Dr. Fiore and the Cessation Subcommittee considered data from Wisconsin and Minnesota that were from fairly brief periods -- six months for Wisconsin and two months for Minnesota -- and, in the case of Wisconsin, only the subpopulation of senior citizens. (U.S. Exhibit 89,470 at VXA 452 0055.) I was provided with information about additional experience in those states.

Q. What were you provided with?

A. I was provided with an annual report prepared by the Wisconsin quitline that includes the all-age caller volumes for 2001 (from May to December), 2002, and 2003 (JD-068086 at 3). I also was provided with an "annual report card review" of the Minnesota quitline prepared by an outside evaluator. The "report card" includes the caller volume for a longer period of time -- fourteen, rather than two months -- after Minnesota made nicotine replacement therapy available through its quitline in September 2002 (JD-068079 at 31).

Q. What did those documents show about the participation rates in Wisconsin and Minnesota?

A. Based on the caller information in those documents, along with state-specific population and smoking prevalence data, the participation rates for Wisconsin and Minnesota are as reflected in the table that is JDEM-060667.

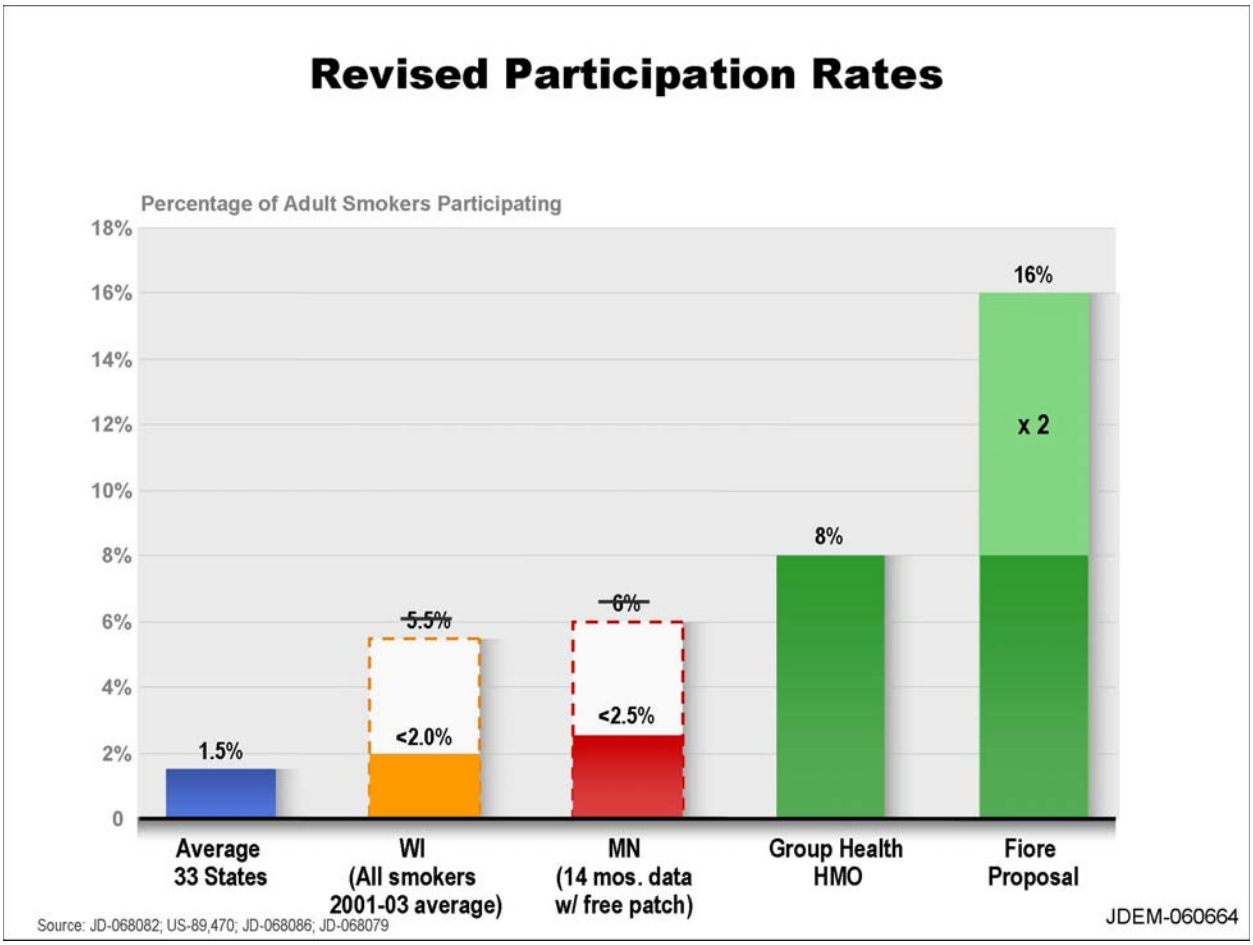
Participation Rate Information

	Wisconsin 2001	Wisconsin 2002	Wisconsin 2003	Minnesota 2003
Census Population	5,405,905	5,440,367	5,474,290	5,064,172
Adult Population Per Group Health Formula*	3,784,134	3,808,257	3,832,003	3,544,920
Smoking Prevalence**	23.6%	23.4%	22.1%	21.1%
Smoking Adults	893,056	891,132	846,873	747,978
Annualized Caller Volume***	17,700	8,560	9,630	17,019
Annual Participation Rate	**1.98%**	**0.96%**	**1.14%**	**2.28%**

*Per U.S 89,470 at VXA 452 0051, line 2 (population times 0.7)
**2001 - MMWR 52(14): 303-307, 304 (2003); 2002 - MMWR 52(53): 1277-1280, 1279 (2004); 2003 - MMWR 53(44): 1035-1037, 1037 (2004)
***JD-068086 at 3 (Wisconsin); JD-068079 at 31 (Minnesota)

JDEM-060667

As reflected there, the annual participation rates for Wisconsin and Minnesota were about 2%. Using those participation rates would change the demonstrative used previously, JDEM-060663, as illustrated in JDEM-060664.



2. OTHER COMPONENTS OF THE NATIONAL CESSATION PROGRAM.

Q. I'd like to turn to Dr. Fiore's proposed $1 billion per year national media campaign. Have you seen evidence the cost of the national media campaign that Dr. Fiore proposes was developed, assessed or evaluated in a statistically valid or reliable manner?

A. No, I haven't.

Q. What about the proposed $500 million per year cessation research program, have you seen any evidence that the cessation research program Dr. Fiore proposes was developed, assessed, or evaluated in statistically valid or reliable manner?

A. No.

Q. And with respect to the $500 million per year clinician training and education program, have you seen any evidence that the clinician training and education program Dr. Fiore proposes was developed, assessed, or evaluated in a statistically valid or reliable manner?

A. No, I haven't.

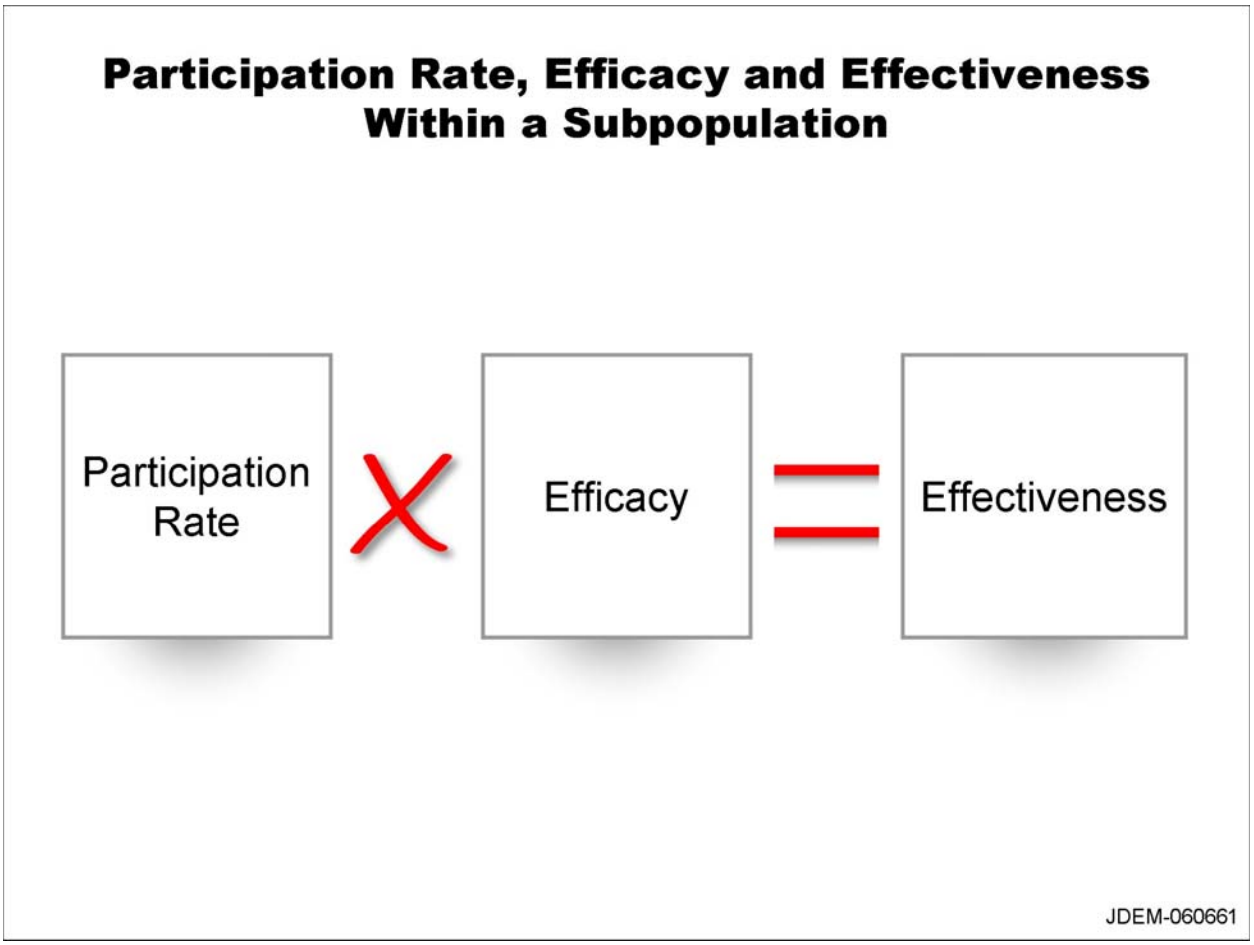
C. THE EFFECTIVENESS OF THE PROPOSED NATIONAL CESSATION PROGRAM.

Q. Now I'd like to turn to Dr. Fiore's estimates of how smokers will fare in terms of quitting if the proposed national quitline is adopted.

A. Fine.

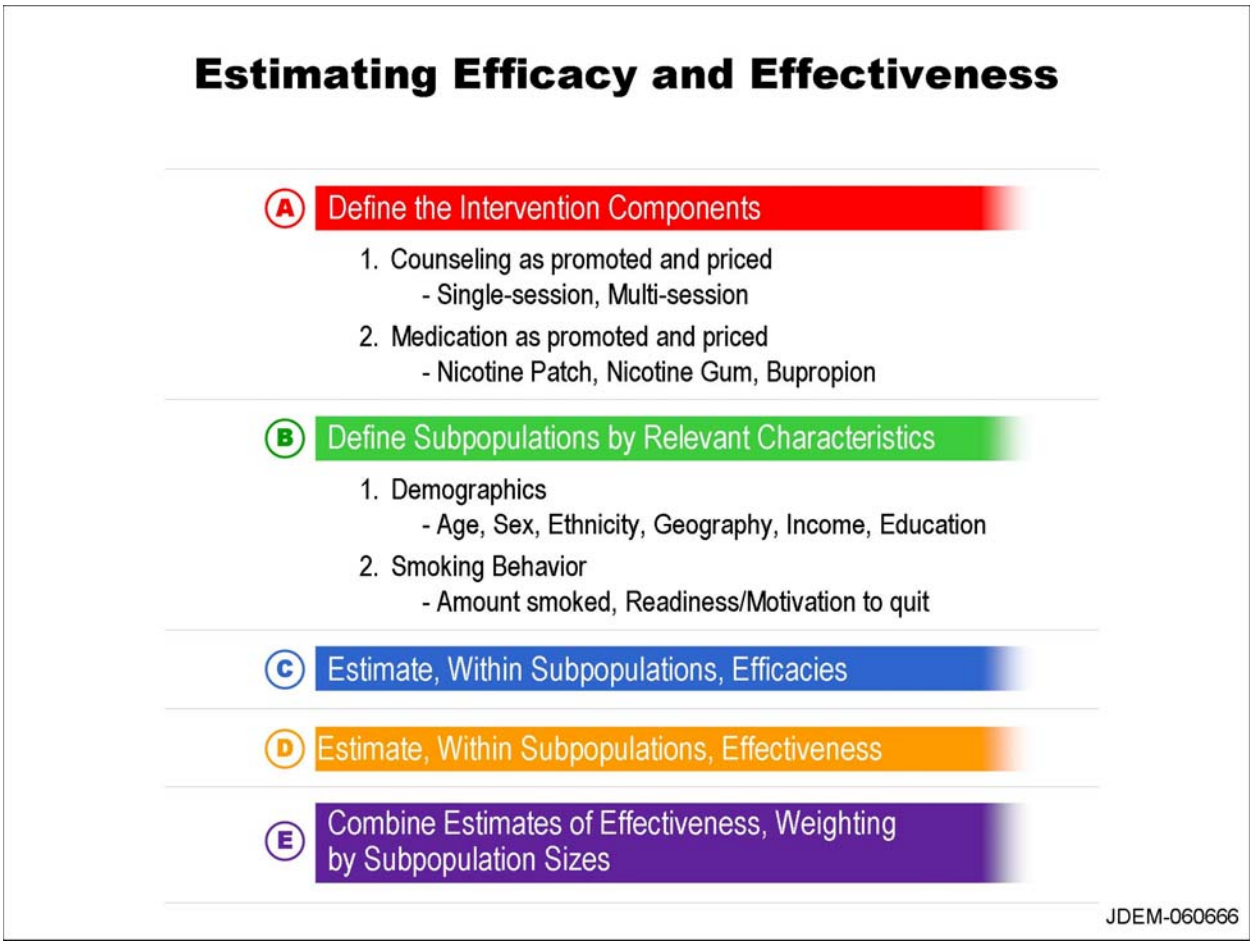
Q. Very generally, how should you go about estimating that?

A. One major factor that we just discussed is how many people use the national quitline -- the participation rate. Another factor that drives the result is the extent to which the smokers who participate are more successful in quitting than they otherwise would have been -- efficacy. And, as reflected in JDEM-060661, the effectiveness of the program in any given subpopulation would be the product of those two factors.



Q. Let's take them one at a time and start with efficacy -- the extent to which smokers who participate are more successful in quitting because of the program. How should one go about estimating efficacy, which is the causal effect, in a statistically valid and reliable manner?

A. The general approach is depicted in JDEM-060666.



As is apparent, the analytical framework generally would be analogous to that used to estimate the participation rate. You would, for each component in the national cessation plan, and for each demographic and smoking behavior subpopulation, estimate efficacy -- the increase in the cessation rate for participants caused by the program.

Q. Why is it important that you look at particular subpopulations?

A. Because cessation interventions do not necessarily have the same efficacy among users in these different subpopulations. (*See, e.g.,* Fiore Written Direct at 61:1 ("certain populations … are not aided by current treatments").)

Q. Would there be problems obtaining sufficiently specific data?

A. Those performing that analysis would have to find that out, but there are hundreds or even thousands of studies that have been done trying to estimate efficacies of various interventions designed to increase smoking cessation. Indeed, Dr. Fiore chaired the group that prepared the 2000 revision of the Clinical Practice Guideline (JD-001210), which stated that it was based on an initial review of 6,000 studies (at iii). So there is a substantial body of existing data, and many of the reported studies may have raw data available with subpopulation-specific information. If the existing data sets have gaps, we can also apply statistical techniques based on explicit assumptions, such as missing data methods, log-linear models, etc., to estimate subpopulation-specific efficacies.

Q. Have you seen any statistically valid or reliable analyses estimating the efficacy of components of the proposed national quitline in subpopulations?

A. I have seen statistically valid and reliable analyses of the efficacy of certain of the components of the proposed national quitline in certain subpopulations, but not for all components or subpopulations. As I just stated, however, there appears to be data available to address at least some of these efficacies.

Q. Dr. Fiore testified that, as set forth in the 2000 Clinical Practice Guideline (JD-001210), each of the individual components of the proposed national cessation program -- telephone-based counseling, nicotine patches, nicotine gum, and bupropion -- was found to increase the rate of smoking cessation in clinical trials. (Fiore Written Direct at 40:20–41:13; 45:7–49:15). Are you disagreeing with the results of those studies?

A. No, I'm not. Although the conclusions reached in some of those studies were, in some instances, overly optimistic in my view, they are among those I was referring to in my prior

answers. Most of those studies were randomized clinical trials estimating the causal effects of those cessation components, and the meta-analyses set forth in the 2000 Clinical Practice Guideline appear to have been done properly.

Q. So isn't that evidence of the efficacy of the components of the national quitline within the different subpopulations?

A. It is evidence that, in the subpopulation of smokers who are sufficiently motivated to enter a clinical smoking cessation randomized trial, the various intervention components had some causal effect on increasing quitting.

Q. What's the difference?

A. Merely because the intervention components have efficacy in subpopulations of persons who were sufficiently ready and motivated to quit to sign up for a clinical trial does not mean that those components will have efficacy in other, different subpopulations. Simply assuming that the efficacies in the subpopulations who participated in comparatively small and carefully monitored clinical trials will apply to all subpopulations recruited by mass media advertising seems unrealistic and is inappropriate. For example, Pierce & Gilpin in "Impact of Over-the-Counter Sales on Effectiveness of Pharmaceutical Aids for Smoking Cessation," *J. Am. Med. Ass'n* 288(10):1260-1264 (2002) (JD-067959), wrote at 1260 that:

> Since becoming available over the counter, NRT [nicotine replacement therapy] appears no longer effective in increasing long-term successful cessation in California smokers.

Similarly, Shiffman et al. in "Real-world efficacy of prescription and over-the-counter nicotine replacement therapy," *Addiction* 97:505-516 (2002) (JD-068092) observed at 506 that:

> The traditional clinical trials approach, in which patients and physicians are prospectively enrolled in a study, does not address real-world practices credibly.

(Citation omitted). Velicer & Prochaska in "An Expert System Intervention For Smoking Cessation," *Patient Educ. & Counseling*, 36(2):119-129 (1999) (JD-061645) noted at 125 that:

> Reactively recruited samples typically involve no more than 5% of the available population and tend to be disproportionately female, highly educated and in the later stages of change.

And Velicer et al. in "Distribution of Smokers by Stage in Three Representative Samples," *Preventive Medicine* 24:401-411 (1995) (JD-046874) stated at 401 that:

> Reactively recruited samples distort estimates of the stage distribution in the population because such samples attract a disproportionate number of late-stage participants….
>
> Existing [smoking cessation] interventions are most appropriate for the Preparation stage, but the majority of the three samples were in the first two stages, resulting in a likely mismatch between the smoker and the intervention.

Q. Can you give an example of why the experience from the subpopulations who participated in randomized clinical trials might not transfer to other subpopulations?

A. Yes. Consider, for example, the efficacy of a nicotine patch in a subpopulation of smokers who were not ready and had little motivation to quit, yet who were using the patch because they saw an advertisement saying it was free and thought that it couldn't hurt. The efficacy of the patch in that subpopulation of smokers would likely be different than its efficacy in a subpopulation of ready-to-quit, motivated-to-quit smokers who signed up for a clinical trial. For example, Fiore et al. in *Clinical Practice Guideline No. 18: Smoking Cessation* (1996) (JD-063828) state at Table 7 on page 41 that "[l]ow motivation" and "[l]ow readiness to change" were among the "[v]ariables associated with lower cessation rates." The smoker's "stage of change" is important for determining what will and will not work as a smoking cessation aid, which is one reason researchers recommend tailoring smoking cessation interventions to specific subpopulations. *See* Prochaska et al., "In Search of How People Change – Applications to

Addictive Behaviors," *Am. Psychol.* 47:1102-1114, at 1106 (1992) (JD-061623). Indeed, the literature suggests that failing to tailor a smoking cessation program to an individual's particular "stage of change" may even have detrimental effects on cessation:

> A person's state of change provides proscriptive as well as prescriptive information on treatments of choice. Action-oriented therapies may be quite effective with individuals who are in the preparation or action stages. These same programs may be ineffective or detrimental, however, with individuals in precontemplation or contemplation stages.

(*Id.*)

Q. Are those potential differences among different subpopulations likely to be of real concern?

A. They certainly could be. As you bring more and more people into a program offering free and heavily-promoted telephone counseling and medications, you are, almost by definition, bringing in additional members of subpopulations of smokers who are less ready and less motivated to quit smoking. If so, then the efficacy rates in those newly-added subpopulations are likely to be lower than the efficacy rates in subpopulations who participated in the clinical trials.

Q. Dr. Fiore testified that the results of studies that involved only a subpopulation of persons who were sufficiently motivated to join a cessation study apply to other, less motivated subpopulations because "odds ratios protect against a result being skewed by self-selection … " (Fiore Written Direct at 47:18-19). Do you agree that the odds ratios obtained from the subpopulations who participated in the clinical trials will apply to all subpopulations?

A. No. Odds ratios are not akin to the Law of Gravity such that they apply equally to all subpopulations. The odds ratio for the cessation components in subpopulations other than the

subpopulations that participated in the clinical trials very likely will differ -- it is simply an empirical issue.

Q. Have you been involved in any instances where researchers raised similar concerns about whether the results obtained in clinical trials are more broadly generalizable?

A. Sure. In one example I mentioned earlier, the General Accounting Office asked me to provide statistical consulting advice on an issue relating to breast cancer treatment. Randomized clinical trials had shown that, in single-center and multi-center cancer centers in the United States and Europe, the 5-year survival rates for mastectomy and lumpectomy were almost the same for a certain category of women with breast cancer. Before using those results as a basis for the population-wide recommendations, however, the GAO wanted to know whether the experience from the subpopulations who participated in the randomized clinical trials in the cancer centers would generalize to the general population. That is, what is the evidence that the results could be "bridged" to the general United States population?

We found that, although there is a better 5-year survival rate for patients in the randomized trials than for those in the general population, there was little difference between the two treatment options being examined in the general population, just as there had been essentially no difference in the randomized trials.

Q. Have you ever been involved in a situation where the results obtained from subpopulations who participated in randomized clinical trials did not apply more generally in the population?

A. Sure. I consulted with a drug company, Roche Pharmaceuticals, in connection with its efforts to market a weight-loss drug, Xenical. Xenical was (and is) a drug taken in connection with dieting to enhance weight loss, and it worked by, in essence, preventing the body from

digesting certain fats in food. By preventing the body from digesting those fats, however, it made some users incontinent if they continued to eat too much fatty food. The drug had proven efficacy in the subpopulations who participated in randomized clinical trials, and its manufacturer was seeking to market it more broadly when I was approached.

Q. What happened?

A. It is my understanding that Xenical was something of a market failure, at least as compared to the manufacturer's hopes based on the number of people who claimed they want to lose weight. Unlike the subpopulations of people who participated in the randomized clinical trials, people in the general population who claimed they wanted to lose weight apparently were not nearly as willing to give up fatty foods to avoid the drug's possible side effect of incontinence.

Q. How is that experience possibly analogous to the experience here?

A. There, the subpopulation of people in the clinical trial were highly motivated and were actually willing to give up fatty foods, and thereby avoid the drug's side effects. Other subpopulations of people who were interested when the drug was marketed more broadly, however, apparently were less motivated or at least less willing to give up fatty food to avoid the drug's side effects. Accordingly, efficacy in the clinical trial did not translate well into effectiveness in the general population because the participation rate was lower than expected based on the number of people claiming that they wanted to lose weight.

As to smoking, the subpopulations of smokers who signed up for clinical trials are, almost by definition, more ready and more motivated to quit smoking than subpopulations of smokers who say they want to quit smoking and might use the various cessation intervention

components if they were heavily advertised and promoted. Accordingly, efficacy in smoking cessation clinical trials may not translate well into effectiveness in the general population.

Q. Let's turn to effectiveness, which you just mentioned. How would you go about estimating the effectiveness of a national quitline such as that proposed by Dr. Fiore in a statistically valid and reliable manner?

A. We would multiply the estimated participation rate by the estimated efficacy rate for each intervention component within each subpopulation. The result would be the estimated effectiveness of that component in that subpopulation. Then we would weight each subpopulation's estimated effectiveness by the relative fraction of that subpopulation in the nation, and then we would add the weighted effectivenesses across the subpopulations to get an estimated effectiveness for each component in the nation. Finally, we would multiply by the number of smokers in the nation to get the estimated number of people who would be expected to quit using that component of the national quitline. This estimation process is outlined in JDEM-060666.

Q. Dr. Fiore testified in his direct testimony that 1 million persons each year will quit smoking using the services offered through the proposed national quitline. (Fiore Written Direct at 70:3-11). Is that its effectiveness?

A. I believe that it is supposed to be an estimate of effectiveness, but Dr. Fiore provided little in the way of analysis, data, or support for that number in his expert disclosure, in his deposition, or in his written direct. He simply states that five million smokers each year will use the quitline and that one million of them will quit. (Fiore Written Direct at 70:3-11; *see also* U.S. Exhibit 89,464 at 14 (Cessation Subcommittee final draft report stating that approximately 1 million smokers would quit each year using the quitline based on a 10% participation rate and a

20% long-term successful cessation rate)). But it is not entirely clear to me that Dr. Fiore's number is supposed to be an estimate of the causal effect of the program.

Q. Are you aware of any statistically valid or reliable basis for Dr. Fiore's conclusion that 1 million smokers would quit smoking each year using the national quitline?

A. No. The 1 million quitters figure appears to be basically an assumption.

Q. Is it your understanding that the 1 million smokers who supposedly will quit using the proposed national quitline are additional quitters who would not quit but for the proposed national quitline?

A. Although Dr. Fiore has not yet testified at trial as I am preparing this written direct, as I said before, his deposition and written direct trial testimony are not clear on this point. During his deposition, Dr. Fiore testified that 1 million smokers would quit smoking each year using the proposed national quitline, but it was not clear to me whether he was saying that (a) this represented 1 million additional quitters as a result of the program (*i.e.*, whose quitting was caused by the program); or (b) the 1 million persons using the proposed national quitline each year were not all additional quitters, but included those who might quit even if there were no program. (*See* Fiore *U.S.* Dep. 5/5/05 at 70:19–76:12). In his written direct testimony, Dr. Fiore testified that his proposed program would "assist[] one million additional smokers to quit every year as a result of the program," but also testified that 1 million smokers would quit each year using the cessation services offered through the program. (Fiore Written Direct at 69:12-13; 70:3-11).

Q. Is there a way one could estimate in a statistically valid or reliable manner the incremental or additional number of smokers who would quit each year as a result of the proposed national quitline?

A. Yes, there is. As I described earlier, we would estimate participation rates and efficacies of each component of the program within subpopulations, and then combine these disaggregated values into one national estimate of the causal effect of the cessation program, as outlined in JDEM-060666.

Q. Have you seen any evidence that Dr. Fiore or the Cessation Subcommittee made a statistically valid or reliable determination of the incremental or additional number of smokers who will quit each year as a result of the proposed national quitline?

A. No, I am aware of no statistically valid or reliable estimation that either Dr. Fiore or the Cessation Subcommittee made of the number of smokers who will quit because of the proposed national quitline (*i.e.*, excluding those who would have quit with or without the proposed national quitline).

Q. In your view, is the incremental or additional number of smokers who would quit each year the measure of effectiveness that is most relevant to determining whether to implement the proposed national quitline?

A. It is. The change in quitting that would result from the proposed national quitline -- its causal effect on smoking cessation -- is the basic measure of effectiveness we should use. Obviously, if a significant percentage of the persons who quit after using the proposed national quitline would have quit without the quitline, then the proposed national quitline does very little to further the goal of increasing cessation.

Q. Before we move on, are you saying that the proposed national cessation program would have no effectiveness -- that no additional smokers would quit smoking if it were adopted?

A. No. What I am saying is that there is no statistically valid or reliable evidence that I have seen presented concerning how many smokers would quit because of the program. At this juncture, nobody has tried to estimate statistically what the proposed national quitline's participation rate or effectiveness would be. Those estimates are the basic information one would need to make a reasoned decision about whether the costs of the program are justified by its benefits.

Q. Dr. Fiore's national cessation program is, as proposed to the Court here, scheduled to last approximately 25 years, do you recall that?

A. Yes, I do.

Q. Have you seen any evidence that the approximately 25-year duration of the national cessation program was developed, assessed or evaluated in a statistically valid or reliable manner?

A. No.

VI. DR. WYANT'S ESTIMATES OF DEATHS, TREATMENT-YEARS-OF-DISEASE, YEARS OF POTENTIAL LIFE LOST, AND HEALTH-CARE COSTS "ATTRIBUTABLE" TO SMOKING ARE NOT STATISTICALLY RELIABLE AND VALID ESTIMATES OF THE CAUSAL EFFECTS OF SMOKING.

Q. Dr. Wyant testified, at 22:15-23 of his written direct testimony, that:

> **Q: Looking at the excerpts in this exhibit, the terms "responsible for," "caused by," "resulting in," and "attributable" are all used. What is the significance of the different terminologies?**
>
> **A: None. As the excerpts indicate, in examining the burden imposed on the United States population by smoking, these terms just provide different ways of saying the same thing. In the context of the scientific work that is summarized in these Surgeon General excerpts, just as in our own calculation for the Youth Addicted Population, phrases such as "caused by smoking," "attributable to smoking," "due to smoking," "smoking related," and "smoking is responsible for" are essentially synonymous.**

Do you agree that "caused by smoking" and "attributable to smoking" are "essentially synonymous?"

A. No, I do not agree.

Q. Why do you disagree?

A. As we discussed, the principles of causal inference to which I have contributed are widely accepted, not only in statistics, but also in economics, psychology, medicine, and other fields. Based on those principles, neither I nor, to my knowledge, other researchers familiar with causal inference, such as Plaintiff's expert, Dr. Zeger, would equate "caused by" and "attributable."

Q. Based on the principles of causal inference to which you have contributed, what is the difference between a smoking "attributable" estimate and an estimate of the causal effect of smoking?

A. Essentially, a smoking "attributable" estimate is merely descriptive. It describes the difference in, say, health-care expenditures among individuals who smoke compared to individuals who have never smoked and who are similar, at best, with respect to those factors that are explicitly included in the attributable risk formula being used, for example, age and gender.

In this case, the smoking-attributable, health-care cost estimates prepared by Dr. Wyant compared total health-care expenditures of individuals with one of their smoking-related diseases to the total health-care expenditures of individuals without one of those diseases, but did not control for other factors that could importantly affect total health-care expenditures, such as insurance status or the occurrence of other medical conditions, such as accidental injuries or mental illness.

To obtain a causal estimate of the effect of smoking on health-care expenditures, however, we have to do something different. We have to compare the health-care expenditures of smokers in the actual world to what their health care expenditures would be if they had not smoked. To do so, one must consider how smokers in a world without smoking would be different, apart from their smoking, not just in one year, but through time, using actual world data and explicit assumptions. For example, it is well-established that individuals who successfully quit smoking usually gain weight. Accordingly, we should include in a causal analysis an assumption that, in a world in which they did not smoke, smokers would usually weigh more than they do in the actual world in which they smoked, suggesting that the appropriate comparison using actual world data would not be to a similar never smoker who has the *same* body mass index as a current smoker, but to a similar never smoker who has a *higher* body mass index.

Q. Apart from your work on causal inference generally, have you commented on precisely these differences between a causal estimate of the effects of smoking and non-causal smoking-"attributable" estimates in your peer-reviewed publications?

A. Yes, the differences between a causal estimate of the effects of smoking and a non-causal, merely descriptive, smoking "attributable" estimate have been addressed in my peer-reviewed publications, specifically my chapter, "Statistical Issues in the Estimation of the Causal Effects of Smoking Due to the Conduct of the Tobacco Industry," in *Statistical Science in the Courtroom* (Gastwirth ed.) pp. 321-351 (2000) (JD-063877) at 328; as well as my article, "Estimating The Causal Effects of Smoking," *Statistics in Medicine*, 20:1395-1414 (2001) (JD-063879).

Q. Were your publications that distinguished estimates of the causal effects of smoking from descriptive, smoking "attributable" estimates, criticized by others knowledgeable about causal inference?

A. No, to the best of my knowledge, I have received no such criticism from anyone. In fact, I gave invited lectures on this precise topic to, among others, the CDC and the American Public Health Association and received no criticism of the work presented.

Q. Have others who are knowledgeable about causal inference embraced the distinction that you draw between causal estimates and smoking-"attributable" estimates?

A. Yes. Dr. Zeger, who submitted two joint expert reports with Dr. Wyant in this case and who drafted portions of their expert report that addressed issues of causal inference (Wyant *U.S.* Trial Tr. 5/2/05, at 20043:17-21), has acknowledged the difference in his own publications and in his deposition in this case.

For instance, in a paper by Dr. Zeger and others, "Disease Cases and Their Medical Costs Attributable to Smoking: An Analysis of the National Medical Expenditure Survey," *J. Econometrics,* 112:135-151 (2003) (U.S. Exhibit 74,081), Dr. Zeger stated (at 139) that "[b]y *attributable*, we imply a comparison of smokers to *otherwise similar* non-smokers. That is, we estimate for the population of people for which the NMES sample is representative, the difference in rates of disease for smokers and similar non-smokers." Dr. Zeger then defines (at 140) what he means by "similar," stating "[b]y *similar*, we mean that individuals have similar values of the covariates including age, gender, race, income level and others available in NMES . . . ."

Dr. Zeger then went on (at 140) to recognize that these were not causal estimates:

> Other investigators (e.g., Rubin, 2001) have discussed estimation of the *causal effects* of smoking, namely the difference in disease

> rates or expenditures for a population of smokers compared to what would have occurred had they never smoked. These counterfactual quantities are not directly observable. Their estimation or extrapolation, is beyond the scope of this paper.

Also, in his deposition in this case (Zeger, *U.S.* Dep., 9/10/02, at 192:2-192:10), Dr. Zeger acknowledged that the smoking-"attributable" estimates he prepared with Dr. Wyant do not answer the causal question of what would have happened if the cohort of so-called "youth-addicted smokers" had never smoked:

> Q. So, to be clear, you are not answering the question of anything having to do with what would have happened if the youth smokers hadn't smoked?
>
> A. I'm not -- that's right. I'm not saying what would have happened to this particular cohort if at some moment in time you wiped out all of their smoking history, they never had smoked.

Q. Thank you, Professor Rubin.