

---

# COMMENTARY

## Validation of Intermediate End Points in Cancer Research

Arthur Schatzkin,\* Laurence S. Freedman, Mark H. Schiffman, Sanford M. Dawsey

---

**Investigations using intermediate end points as cancer surrogates are quicker, smaller, and less expensive than studies that use malignancy as the end point. We present a strategy for determining whether a given biomarker is a valid intermediate end point between an exposure and incidence of cancer. Candidate intermediate end points may be selected from case series, ecologic studies, and animal experiments. Prospective cohort and sometimes case-control studies may be used to quantify the intermediate end point-cancer association. The most appropriate measure of this association is the attributable proportion. The intermediate end point is a valid cancer surrogate if the attributable proportion is close to 1.0, but not if it is close to 0. Usually, the attributable proportion is close to neither 1.0 nor 0; in this case, valid surrogacy requires that the intermediate end point mediate an established exposure-cancer relation. This would in turn imply that the exposure effect would vanish if adjusted for the intermediate end point. We discuss the relative advantages of intervention and observational studies for the validation of intermediate end points. This validation strategy also may be applied to intermediate end points for adverse reproductive outcomes and chronic diseases other than cancer.** [J Natl Cancer Inst 82:1746-1752, 1990]

---

The biologic events in the human body that take place between carcinogen exposure and the development of a malignancy can be viewed as "intermediate end points" in carcinogenesis. These end points may be discrete events such as formation of a large-bowel adenomatous polyp, or they may be quantitative changes such as an increase in the proliferation rate of epithelial cells.

Because malignancies develop slowly and relatively infrequently, investigations of the relation between suspected carcinogenic exposures and cancer may require many participants observed over a long period of time. Intermediate end points, however, can be assessed earlier than malignancy. Moreover, if discrete events, the intermediate end points are

usually less rare than the cancer end point, thus permitting quicker, smaller, and less costly investigations. For example, a study of diet in relation to serum estradiol (1), or proliferation of epithelial cells of the large bowel (2), could be carried out on several dozen subjects in a few months, whereas a dietary intervention study with breast or large-bowel cancer as end points would require tens of thousands of subjects with follow-up in excess of 5 years.

An underlying assumption of many studies using biomarkers as end points is that any observed relation between an exposure and the marker will translate into a similar relation between exposure and cancer per se. We will address this crucial assumption and present a strategy for validating biomarkers as true intermediate end points. Our purpose in outlining these methods is to stimulate studies that provide data for validating cancer intermediate end points.

To simplify our discussion, we will assume that a biomarker of interest (a potential intermediate end point) is a discrete event. The arguments we use can be extended to quantitative markers, but we do not address this issue here.

### Identifying Potential Intermediate End Points

In epidemiologic research, case series and ecologic (correlational) studies are considered to be "hypothesis generating" with regard to the relation between various exposures and diseases. Animal experiments also may yield exposure-disease hypotheses that merit further investigation in humans. Hypotheses are then investigated more definitively by observational studies (case-control and cohort) of individuals and may ultimately be tested in intervention studies (3). A similar broad research strategy may be carried out first to identify

---

Received July 16, 1990; revised September 17, 1990; accepted September 20, 1990.  
A. Schatzkin, L. S. Freedman, S. M. Dawsey (Division of Cancer Prevention and Control), M. H. Schiffman (Division of Cancer Etiology), National Cancer Institute, Bethesda, Md.

\*Correspondence to: Arthur Schatzkin, M.D., Dr.P.H., Executive Plaza North, Room 211, National Institutes of Health, Bethesda, MD 20892.

biomarker candidates and then to validate these markers as intermediate end points.

To select potential intermediate end point candidates from the large pool of available biomarkers, we can use case series, ecologic studies, or animal experiments. In a case series, we may find a specific marker in the overwhelming majority of persons with cancer. Several studies have demonstrated epithelial cell hyperproliferation in persons with large-bowel cancer or adenomas (2). Human papillomavirus DNA has been found in cervical cell samples from 80% to 90% of women with invasive cervical cancer (4). In an ecologic study, we measure the prevalence of a biomarker in populations at varying risks of cancer. As examples, both fecal mutagenicity (5) and cell proliferation indexes (6) have been found to be higher in populations at greater, as opposed to lesser, risk of large-bowel cancer. Animal experiments have suggested a number of potential intermediate end points. Increased epithelial cell proliferation, aberrant crypt production, and microadenoma formation have been observed in rodents administered potent chemical carcinogens (7).

## Beginning the Validation Process: Confirming the Biomarker–Cancer Link

### Case–Control Studies

Once we have identified a potentially interesting biomarker through case series, ecologic studies, or animal experiments, a logical progression in research strategy is to conduct a case–control study, in which the prevalence of the biomarker in cases is compared with that in controls.

A persistent concern with case–control studies of biomarkers is the possibility of reverse causality, whereby the disease affects marker determinations and creates a misleading case–control difference. Fecal bile acid concentration, for example, has been proposed as a relevant biomarker of large-bowel carcinogenesis, but measurements of fecal bile acids in large-bowel cancer cases may be altered beyond interpretation by bleeding in the large bowel and changes in bowel habits that characterize this malignancy (Schiffman M: unpublished data).

Case–control studies, however, may be quite informative in situations where the possibility of reverse causality is unlikely, eg, the relation between human papillomavirus infection and cervical cancer. Although it is possible that elevated prevalence of human papillomavirus infection in women with cervical cancer is a consequence of the disease, the strong oncogenic potential of the virus *in vitro* argues against this possibility. For this candidate intermediate end point, case–control studies are potentially valuable.

### Cohort Studies

Prospective cohort studies of potential intermediate end points avoid the reverse causality problem because intermediate end point status is assessed prior to the development of cancer. In this type of study, the incidence of cancer is compared between persons who are intermediate end point positive and those who are intermediate end point negative. As an example, M. Wargovich (personal communication) has initiated a prospective cohort study of cell proliferation in

relation to subsequent adenomatous polyp recurrence. (In this case, the outcome is neoplasia—adenomatous polyps—rather than cancer.) This study will determine initially the labeling index for each person undergoing polypectomy and observe whether that person has a subsequent polyp recurrence.

## Quantifying the Marker–Cancer Association

### Attributable Proportion

We now wish to use our case–control and cohort data to quantify the relation between a putative intermediate end point and cancer. The epidemiologic measure of association most appropriate for this purpose is the *attributable proportion* (AP), which can be defined as the proportion of cases of disease that is attributable to the intermediate end point. AP is determined directly from two other concepts in the epidemiology and screening literature, *sensitivity* and *relative risk* (3).

Case–control, cohort, or intervention studies can provide data on sensitivity and relative risk. The sensitivity is simply the proportion of patients with cancer in the case group who are intermediate end point positive, ie,  $A/(A + C)$  (Table 1). The relative risk is the incidence of cancer in those who are intermediate end point positive divided by the incidence of cancer in those who are intermediate end point negative, ie,  $[A/(A + B)]/[C/(C + D)]$  (Table 1).

The formula for AP can be written as

$$AP = S(1 - 1/R), \quad [1]$$

where  $S$  is the sensitivity and  $R$  is the relative risk (3). (We assume here that follow-up time is long enough for AP to be stable with respect to duration of follow-up.) Values of AP for different values of  $S$  and  $R$  are shown in Table 2. The table shows that  $S$  tends to be more important than  $R$  in determining AP. If  $S$  is low, even a large  $R$  will result in only a relatively low AP. Conversely, a modest  $R$  yields a relatively high AP when  $S$  is high.

The AP allows us to quantify the change in cancer incidence in relation to a given change in the proportion of persons who are intermediate end point positive, assuming that intermediate and cancer end points are causally linked. If the intermediate end point positivity were eliminated from the population, we would expect the cancer incidence to be reduced by  $AP \times 100\%$ . In general, if  $D_{IE}$  is the percent reduction in the proportion of persons who are intermediate end point positive, then the resulting percent reduction in cancer incidence,  $D_{CA}$ , is given by (8)

$$D_{CA} = D_{IE} \cdot AP. \quad [2]$$

### Inferences From the Attributable Proportion

The nature of the underlying causal pathway for an intermediate end point has considerable bearing on the AP.

Table 1. General relation of intermediate end points and cancer

		Cancer	
		Yes	No
Intermediate end point	+	A	B
	–	C	D

**Figure 1.** Single intermediate end point ( $IE_1$ ) is necessary for cancer (CA).



Fig 1 depicts a hypothetical causal relation between a single intermediate end point ( $IE_1$ ) and cancer, according to which  $IE_1$  is *necessary* for the development of cancer. In this case  $S = 1.0$ ,  $R = \text{infinity}$  (in Table 1,  $C = 0$ ), and  $AP = 1.0$ . An AP of 1.0 implies that the carcinogenic process always works through  $IE_1$ , and one may safely use that intermediate end point as a cancer surrogate.

In reality, one can expect more than one pathway to exist (Fig 2). In this case, some cancers develop through  $IE_1$ , but the intermediate end point  $IE_1$  is *not necessary* for cancer to occur. Because some cancers may be causally preceded by  $IE_2$ , and not by  $IE_1$ , the  $S$  of that intermediate end point ( $IE_1$ ) is less than 1.0. If  $S$  is less than 1.0, then the AP for  $IE_1$  is also less than 1.0 (equation 1) and  $D_{CA}$  is less than  $D_{IE}$  (equation 2). Therefore, reducing intermediate end point positivity by a given percentage will reduce cancer incidence by a lesser percentage.

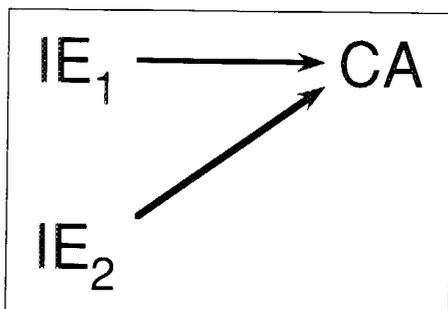
Just as an AP of 1.0 signifies that an intermediate end point will make a valid cancer surrogate, so an AP of 0 for  $IE_1$  implies that  $IE_1$  will not be a valid cancer surrogate. If the AP for  $IE_1$  is 0, then the  $IE_1 \rightarrow CA$  pathway (Fig 2) does not exist and the carcinogenic process operates only through  $IE_2$ .

We give two examples here. Recent studies of human papillomavirus infection and cervical cancer (4) have shown that  $S = 0.9$  and that  $R = 10$  (or greater). AP is then calculated from equation 1 as approximately 0.8, ie,  $0.9 \times [1 - (1/10)]$ . This implies that an intervention that eliminated 30% of human papillomavirus infection would decrease cancer incidence by  $30\% \times 0.8 = 24\%$ , assuming a causal relation.

Fecal mutagenicity has been suggested as an intermediate end point in large-bowel carcinogenesis (5). In a recent case-control study of large-bowel cancer, investigators observed a particular type of fecal mutagenicity in 12% of cases with an estimated  $R$  of 4.4 (9). The calculated AP for this type of mutagenicity is 0.09, ie,  $0.12 \times [1 - (1/4.4)]$ . This implies that an intervention diminishing mutagenicity by 30% would reduce the incidence of large-bowel cancer by only 2.7% ( $30\% \times 0.09$ ).

### More Definitive Validation: Is the Marker a Mediator?

Given that we are likely to encounter case-control or cohort studies yielding APs that are close to neither 0 nor 1.0, what



**Figure 2.** Two intermediate end points ( $IE_1$  and  $IE_2$ ) lead to cancer (CA).

other information can be used to decide whether a biomarker with an AP in the middle range is a valid cancer surrogate? Our key idea is to consider as a whole the relations among exposure, intermediate end point, and cancer. In particular, we want to study whether the given intermediate end point can be shown to mediate an established exposure-cancer relation (10). This can be done in intervention or observational studies.

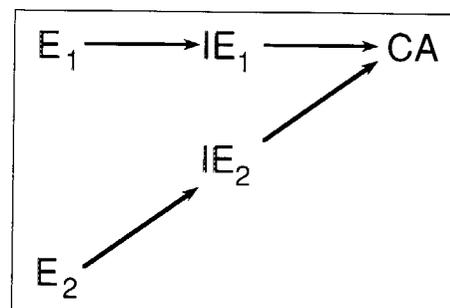
### Intervention Studies

The question of mediation can be particularly well addressed by integrating the assessment of the putative intermediate end point into an intervention study. Prentice (11) has discussed this issue in the context of surrogate end points for clinical trials. The relevant question becomes the following: to what extent does a change in the intermediate end point account for any observed intervention effect? If the intervention is randomly assigned, then the question may be assessed without concern over spurious associations between the intervention and intermediate end point or between the intervention and cancer.

In an intervention study, one has the following information for each participant: intervention group assignment ("exposure"), intermediate end point (measured at a period of time after the intervention has begun), and cancer outcome (yes or no). Let us assume that the intervention "works," that cancer incidence in the intervention group is significantly reduced compared with that in the control group. Since individual data on the intermediate end points are available, it is possible to determine statistically, through stratified adjustment or regression techniques, the extent to which the reduction in incidence is explained by a reduction in intermediate end point positivity.

Fig 3 shows the causal pathways depicted in Fig 2 but with the exposure link added to complete the exposure-intermediate end point-cancer pathway. The AP for  $IE_1$  is between 0 and 1.0, since some, but not all, cancers will be preceded causally by  $IE_2$ . In Fig 3, the effect of the intervention ( $E_1$ ) is fully mediated by its associated intermediate end point ( $IE_1$ ), so we expect the "intermediate end point-adjusted" intervention effect to become zero.

In the hypothetical data shown in Table 3, one sees that in *both* the intervention and control groups the cancer occurrence proportion is five times greater among those who are marker positive than among those who are marker negative. The 33% [ie,  $(36 - 24)/36 \times 100\%$ ] reduction in cancer occurrence with intervention is completely accounted for by the halving of marker positivity with intervention (50% and 25% marker positive, respectively, in the control and



**Figure 3.** Two separate exposures ( $E_1$  and  $E_2$ ), each fully mediated by a different intermediate end point, ( $IE_1$  and  $IE_2$ , respectively), lead to cancer (CA).

Table 2. Attributable proportion [AP = S(1 - 1/R)]

Sensitivity (S) of intermediate end point	Relative risk (R) for intermediate end point positives									
	1.5	2.5	3.5	4.5	5.5	6.5	7.5	8.5	9.5	10.5
0.1	0.03	0.06	0.07	0.08	0.08	0.08	0.09	0.09	0.09	0.09
0.3	0.10	0.18	0.21	0.23	0.25	0.25	0.26	0.26	0.27	0.27
0.5	0.17	0.30	0.36	0.39	0.41	0.42	0.43	0.44	0.45	0.45
0.7	0.23	0.42	0.50	0.54	0.57	0.59	0.61	0.62	0.63	0.63
0.9	0.30	0.54	0.64	0.70	0.74	0.76	0.78	0.79	0.81	0.81

intervention groups). To see this, note that the marker-positive study participants have the same cancer proportion (60%) whether in the intervention or in the control group. The same is true of marker-negative participants, who have a 12% cancer proportion in both study groups. Note that the 33% reduction in cancer proportion satisfies equation 2, where  $D_{CA} = 33\%$ , the control group AP = 0.67, and  $D_{IE} = 50\%$ .

If one were to adjust the intervention effect for the difference in the proportions of marker-positive participants in the two study groups, either by standardization or by regression of the cancer proportion on variables indicating intervention assignment and marker status, the intervention effect would be eliminated. In other words, the intervention works entirely through the intermediate end point. We could then conclude that the intermediate end point  $IE_1$  is a valid cancer surrogate for studies involving  $E_1$  (but *not* necessarily for studies involving other types of intervention).

Fig 4, however, depicts a different situation. Here the effect of the intervention ( $E_1$ ) is only *partially* mediated by the intermediate end point ( $IE_1$ ), since  $E_1$  affects cancer incidence through  $IE_2$  as well as  $IE_1$ . In this case, the  $IE_1$ -adjusted effect of intervention  $E_1$  would no longer be equal to zero but would represent the average effect of  $E_1$  on cancer mediated through an alternative intermediate end point ( $IE_2$ ). Hence, the larger the adjusted intervention effect, the greater is the role played by the alternative pathway. Clearly, it would be hazardous to use  $IE_1$  as a cancer surrogate for assessing other interventions with a mode of action similar to  $E_1$  unless the adjusted treatment effect is quite close to zero. Other interventions, even if they decreased  $IE_1$ , could possibly *increase*  $IE_2$  and thereby offset the cancer reduction achieved through the effect on  $IE_1$ . Without explicitly observing the cancer end point or identifying the other pathway ( $IE_2$ ), we would remain ignorant of this effect. Note that for similar reasons equation 2 does not hold in this setting.

In the case where  $IE_1$  accounted for none of the treatment effect (the marker is not on the  $E \rightarrow CA$  pathway), the marker-adjusted intervention effect would be equal to the unadjusted intervention effect. In this instance, the marker would definitely not be a valid surrogate end point for cancer.

In practice, because of statistical fluctuations, it would be most unusual for the numbers to fall out as perfectly as in Table 3, even when the intermediate end point does entirely mediate the intervention effect. One would, however, calculate the unadjusted and adjusted intervention effect: if the unadjusted effect were large and significant and the adjusted effect were much smaller and nonsignificant, this would provide supporting evidence for the mediating effect of the intermediate end point. If, however, the unadjusted intervention effect were small and nonsignificant, then there

would be no opportunity of positively validating the intermediate end point, although the demonstration of a simultaneous large and significant effect of the intervention on the intermediate end point would suggest that this marker should not be used as a cancer surrogate.

### Observational Studies

Observational studies could also be suitable for studying whether the marker mediates the exposure effect. Considerations similar to those discussed for an intervention study hold for an observational study, but information on exposure/nonexposure is substituted for intervention/nonintervention assignment. When intermediate end points are validated in intervention or observational studies, the inherent temporal assumption is that the exposure precedes the measurement of the intermediate end point, which, in turn, precedes the disease outcome. In an intervention study, the investigation begins at the same time that exposure is first measured, ie, when the treatment group is assigned.

In an observational study, on the other hand, one may be able to decrease the requirement for follow-up time by assessing exposure retrospectively, at the time the intermediate end point is measured. In a cohort study, both exposure and intermediate end point may be assessed prior to the diagnosis of cancer. In a case-control study, however, both exposure and intermediate end point are measured after cancer diagnosis. Thus, observational validation studies (both cohort and case-control) operate on the assumption that exposure precedes intermediate end point. Furthermore, whereas intermediate end point clearly precedes diagnosis of cancer in the cohort study, the possibility that cancer has altered intermediate end point (reverse causation) remains a concern in the case-control approach.

As an example of an observational validation study, investigators at the National Cancer Institute recently established a cohort study of human papillomavirus infection of the cervix, with measurement of the infection at enrollment. The study was to determine whether human papillomavirus

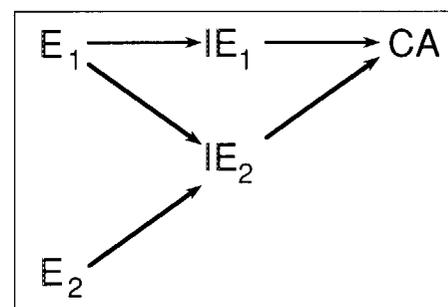


Figure 4. Two separate exposures ( $E_1$  and  $E_2$ ) lead to cancer (CA).  $E_1$  is mediated by two different intermediate end points ( $IE_1$  and  $IE_2$ ).  $E_2$  is fully mediated by  $IE_2$ .

**Table 3.** Hypothetical data from intervention study incorporating an intermediate end point

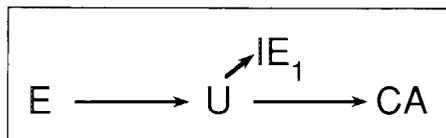
Group	No. of participants	Cancer proportion, %*
Intervention		
Marker positive	25	(15/25) = 60
Marker negative	75	(9/75) = 12
Total	100	(24/100) = 24
Control		
Marker positive	50	(30/50) = 60
Marker negative	50	(6/50) = 12
Total	100	(36/100) = 36

\*Values in parentheses = No. of participants marker positive or marker negative/total No. of participants.

infection mediates the risk of cervical neoplasia associated with the traditional epidemiologic variables "lifetime number of sexual partners" and "age at first sexual intercourse." In this study, the sexual variables can be viewed as prior exposures that can be assessed reasonably accurately at enrollment. The expectation is that the risk of cervical neoplasia will be increased among women who initiated sexual intercourse at earlier ages and who have had greater numbers of sexual partners. If human papillomavirus infection is the intermediate end point explaining these associations, then the excess risk for the sexual variables should disappear; ie, the human papillomavirus infection-adjusted relative risk for the sexual variables should approximate 1.0.

As illustrated above, the observational study has three important advantages over the intervention study with regard to validation of intermediate end points: 1) prior exposures may be ascertained at the time of the intermediate end point assessment, thereby reducing the overall follow-up time; 2) the exposure chosen for study may be a risk factor that has a well-established association with cancer, whereas the intervention may turn out not to modify cancer incidence; and 3) several exposures may be studied simultaneously and the intermediate end point validated against each. However, in observational studies, spurious correlation between the exposure and intermediate end point or disease may exist even after attempts at statistical adjustment. Moreover, implicit assumptions about temporal sequence (between exposure and intermediate end point in both cohort and case-control studies and between intermediate end point and cancer for the case-control investigation) render the observational study inferentially somewhat weaker than the intervention study for intermediate end point validation. In practice, both intervention and observational studies are likely to be useful, and opportunities for validation in observational studies may present themselves more frequently.

**Figure 5.** An exposure (E) leading to cancer (CA) operates through an unobservable (U) event that in turn leads to an intermediate end point (IE<sub>1</sub>) that is not directly on the cancer pathway.



## Further Considerations

### Necessity and Sufficiency in Intermediate End Point Pathways

The causal pathways as they are drawn in Figs 2 to 4 imply that the exposures and intermediate end points are each sufficient for the carcinogenic process to take place. This is an oversimplification of actual causal processes in carcinogenesis; the previous arguments in this commentary do not depend on any such assumption. Not every person on, say, a high-fat diet develops adenomatous polyps, nor does every polyp go on to malignancy. This notion of causal sufficiency and insufficiency in carcinogenesis is reflected in the following pathway:

$$E_1 (+ E_2 + \dots) \rightarrow IE_1 (+ E_3 + \dots) \rightarrow CA.$$

If additional exposures, indicated as (+ E<sub>2</sub> + ...), are necessary in the E<sub>1</sub> → IE<sub>1</sub> pathway, then E<sub>1</sub> by itself would not lead to polyp formation (IE<sub>1</sub>). (E<sub>2</sub> + ...) may represent certain susceptibility factors (12) or other dietary factors such as low vegetable or dietary fiber (13) intake.

The IE<sub>1</sub> (+ E<sub>3</sub> + ...) → CA part of the pathway just described indicates that other exposures may be necessary for a polyp to develop into cancer. If these additional exposures are not present, then the polyp does not go on to cancer. Note that these additional exposures (E<sub>3</sub> + ...) might be exogenous environmental factors such as diet, or they could be endogenous metabolic phenomena such as pH (14) or ratio of secondary to primary bile acids (15) that may be considered intermediate end points in their own right.

### Closely Linked "Intermediates"

In Fig 5 the intermediate end point IE<sub>1</sub> is not directly on the causal pathway from exposure to cancer. Instead this intermediate end point is linked to cancer by means of an unobservable event (U) that leads to both the intermediate end point and cancer. For the intermediate end point to be a useful cancer surrogate in this situation, the time from U to the intermediate end point would have to be shorter than the time from U to cancer.

When an intervention works at the E → U portion of the causal pathway (Fig 5), the effect of the intervention on cancer will be reflected by the effect of the intervention on the intermediate end point. However, it is theoretically possible that the intervention could operate at U → IE<sub>1</sub>, but not at U → CA. In that case, the intervention would affect the intermediate end point but would not have a comparable impact on cancer. Conversely, if the intervention operates at U → CA but not at U → IE<sub>1</sub>, then the intervention would have an effect on cancer that would not be reflected in an intermediate end point effect.

Unfortunately, the situation in Fig 5 cannot be distinguished from that in Fig 4 by estimation of the intermediate end point-adjusted intervention (or exposure) effect. In both cases, this effect will be nonzero. For the pathway in Fig 5, the adjusted effect will lie between zero and the unadjusted effect, and its size will depend on the (unobservable) correlation between IE<sub>1</sub> and U. If the correlation is high, then the adjusted intervention effect will be close to zero; if the correlation is low, then the adjusted effect will be closer to the unadjusted effect. Other

evidence is required to distinguish between the two possible models reflected in Figs 4 and 5.

Micronuclei, which have been identified in several epithelial tissues including those of the mouth, esophagus, lung, and large bowel, have been proposed as a useful intermediate end point in cancer studies (16) and could serve as an example of IE<sub>1</sub> in Fig 5. A micronucleated cell is usually not viable and therefore cannot be a direct cellular precursor of a malignancy. However, the presence of micronucleated cells may reflect an unobservable genotoxic event (U in Fig 5) that may be a necessary step on the pathway to cancer.

### Measurement Error

In the description above, we have assumed that the intermediate end point is measured without error. The presence of measurement error will distort the relations we have described. For example, suppose the intermediate end point lies on the unique causal pathway between an exposure and disease (Fig 3). The intermediate end point-adjusted exposure effect, which is truly zero, will tend to be estimated as a nonzero quantity as a result of the measurement error. This situation is analogous to the effect on risk estimation of misclassification of confounding variables (17). When examining such adjusted exposure effects, we may thus confuse the effects of measurement error with the models in Figs 4 and 5. In practice, measurement error will always be present, but knowledge of the error variance allows one to correct for its effect. Statistical methods for such corrections may be found in Fuller (18), and their importance is beginning to be widely appreciated by epidemiologists (19).

### Sample Size

Validating an intermediate end point within a cohort or intervention study need not lead to any increase in the required sample size. First, we would generally expect the association between the intermediate end point and cancer to be at least as strong as the association between exposure and cancer. Therefore, if the study is sufficiently large to detect the effect of exposure, it will also be large enough to detect the relation of the intermediate end point with cancer. Second, the estimate of the intermediate end point-adjusted exposure effect will usually have only marginally less precision than the estimate of the unadjusted exposure effect. The adjusted exposure effect can therefore be studied with almost the same statistical power as the adjusted treatment effect. Further detailed work on the sample size requirements for a validation study is available from the authors.

### Conclusion

In spite of the undeniable attractiveness of studies using intermediate end points as surrogates for cancer, such studies do not obviate the need for large-scale prospective investigations. Intermediate end points have to be validated if they are to be used reasonably as cancer surrogates.

We have not yet had the opportunity to try out the validation method we propose on data relating to cancer; such data may not be available for some time. Our arguments,

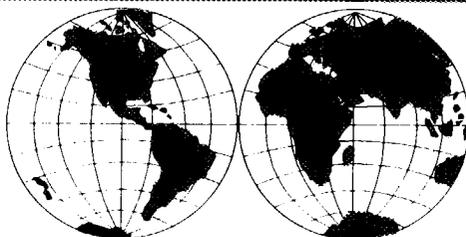
however, are also applicable to the validation of intermediate end points for other diseases. Data from cardiovascular disease prevention trials have become available for validating serum cholesterol as an intermediate end point for coronary heart disease (20,21). Validation studies of semen biomarkers (22) as intermediate end points for adverse reproductive outcomes such as low birth weight and congenital malformations have the decided advantage of a relatively quick occurrence of the final end point. Studies of CD4 counts in acquired immunodeficiency syndrome (23) or pulmonary function tests in neoplastic and nonneoplastic lung disease (24,25) present other potential opportunities for the application of this validation approach.

In most cases, the validation strategy discussed requires the incorporation of intermediate end points in ongoing and new cohort and intervention studies. This will be an expensive and time-consuming process that is just beginning at the present time. We cannot stress too strongly, however, that these validation studies are essential if intermediate end points are to progress beyond being merely interesting phenomena that may or may not tell us something about cancer and its causes.

### References

- (1) PRENTICE R, THOMPSON D, CLIFFORD C, ET AL: Dietary fat reduction and plasma estradiol concentration in healthy postmenopausal women. *J Natl Cancer Inst* 82:129-134, 1990
- (2) LIPKIN M: Biomarkers of increased susceptibility to gastrointestinal cancer: New applications to studies of cancer prevention in human subjects. *Cancer Res* 48:235-245, 1988
- (3) ROTHMAN KJ: *Modern Epidemiology*. Boston: Little, Brown, 1986
- (4) LORINCZ AT, TEMPLE GF, KURMAN RJ, ET AL: Oncogenic association of specific human papillomavirus types with cervical neoplasia. *JNCI* 79:671-677, 1987
- (5) SCHIFFMAN MH: Epidemiology of fecal mutagenicity. *Epidemiol Rev* 8:92-105, 1986
- (6) LIPKIN M, UEHARA K, WINAWER S, ET AL: Seventh-Day Adventist vegetarians have a quiescent proliferative activity in colonic mucosa. *Cancer Lett* 26:139-144, 1985
- (7) COOKE T, KIRKHAM N, STAINTHORP DH, ET AL: Detection of early neoplastic changes in experimentally induced colorectal cancer using scanning electron microscopy and cell kinetic studies. *Gut* 25:748-755, 1984
- (8) STURMANS F, MULDER PGH, VALKENBURG HA: Estimation of the possible effect of interventive measures in the area of ischemic heart diseases by the attributable risk percentage. *Am J Epidemiol* 105:281-289, 1977
- (9) SCHIFFMAN MH, ANDREWS AW, VAN TASSELL RL, ET AL: Case-control study of colorectal cancer and fecal mutagenicity. *Cancer Res* 49:3420-3424, 1989
- (10) SUSSER M: *Causal Thinking in the Health Sciences*. New York: Oxford Univ Press, 1973
- (11) PRENTICE RL: Surrogate endpoints in clinical trials: Definitions and operational criteria. *Stat Med* 8:431-440, 1989
- (12) CANNON-ALBRIGHT LA, SKOLNICK MH, BISHOP DT, ET AL: Common inheritance of susceptibility to colonic adenomatous polyps and associated colorectal cancers. *N Engl J Med* 319:533-537, 1988
- (13) TROCK B, LANZA E, GREENWALD P: Dietary fiber, vegetables, and colon cancer: Critical review and meta-analyses of the epidemiologic evidence. *J Natl Cancer Inst* 82:650-661, 1990
- (14) BRUCE WR: Recent hypotheses for the origin of colon cancer. *Cancer Res* 47:4237-4242, 1987
- (15) HILL MJ: Mechanisms of colorectal carcinogenesis. In *Diet and Human Carcinogenesis: Proceedings of the 3rd Annual Symposium of the European Organization for Cooperation in Cancer Prevention Studies (ECP)*, Aarhus, Denmark, June 19-21, 1985 (Joossens JV, Hill MJ, Geboers J, eds). Amsterdam: Elsevier, 1985, pp 149-163
- (16) STICH HF, STICH W, ROSIN MP, ET AL: Use of the micronucleus test to monitor the effect of vitamin A, beta-carotene and canthaxanthin on the buccal mucosa of betel nut/tobacco chewers. *Int J Cancer* 34:745-750, 1984

- (17) GREENLAND S, ROBINS JM: Confounding and misclassification. *Am J Epidemiol* 122:495-506, 1985
- (18) FULLER WA: *Measurement Error Models*. New York: Wiley, 1987
- (19) BYAR DP, GAIL MH: Errors-in-variables workshop. *Stat Med* 8:1027-1029, 1989
- (20) LIPID RESEARCH CLINICS PROGRAM: The Lipid Research Clinics Coronary Primary Prevention Trial results. I. Reduction in incidence of coronary heart disease. *JAMA* 251:351-364, 1984
- (21) MANNINEN V, ELO MO, FRICK MH, ET AL: Lipid alterations and decline in the incidence of coronary heart disease in the Helsinki Heart Study. *JAMA* 260:641-651, 1988
- (22) STILLMAN RJ, ROSENBERG MJ, SACHS BP: Smoking and reproduction. *Fertil Steril* 46:545-566, 1986
- (23) GOEDERT JJ, KESSLER CM, ALEDORT LM, ET AL: A prospective study of human immunodeficiency virus type 1 infection and the development of AIDS in subjects with hemophilia. *N Engl J Med* 321:1141-1148, 1989
- (24) COHEN BH: Is pulmonary dysfunction the common denominator for the multiple effects of cigarette smoking? *Lancet* 2:1024-1027, 1978
- (25) TOCKMAN MS, ANTHONISEN NR, WRIGHT EC, ET AL: Airways obstruction and the risk for lung cancer. *Ann Intern Med* 106:512-518, 1987



## How in the world do you find out about the earth?

The Earth Science Data Directory (ESDD) is compiled and produced by the US Geological Survey, an agency of the Department of the Interior and the Federal Government's largest earth-science research agency.

References in the ESDD include information about:

- Data bases concerned with the geologic, hydrologic, cartographic, and biologic sciences
- Data that supports the protection and management of natural resources
- Geographic, sociologic, economic, and demographic data sets

To secure information about becoming an ESDD user, write or call:

ESDD Project Manager  
U.S. Geological Survey  
801 National Center  
Reston, Virginia 22092  
(703) 648-7112 or FTS 959-7112