**Point/Counterpoint**

Point: Surrogate End Point Biomarkers Are Likely To Be Limited in Their Usefulness in the Development of Cancer Chemoprevention Agents against Sporadic Cancers

---

**William B. Armstrong, Thomas H. Taylor, and Frank L. Meyskens, Jr.**

Departments of Otolaryngology [W. B. A.] and Medicine [T. H. T., F. L. M.], Chao Family Comprehensive Cancer Center, University of California, Irvine, Orange, California 92868, and Surgical Service, Veterans Affairs Medical Center-Long Beach, Long Beach, California 90822 [W. B. A.]

**Introduction**

The notion that SEBMs\(^3\) can predict the effectiveness of chemopreventive agents against carcinogenesis in humans has motivated tremendous efforts to development of these compounds (1–3). Successful development of useful SEBMs in AIDS and cardiovascular disease seems to confirm the concept. Why should it not also be valid for chemoprevention of cancer (2)? Unfortunately, oncology has yet to replicate the success of the AIDS and cardiovascular communities in developing reliable SEBMs, but not for want of trying. Although chemopreventive agents have demonstrated some success in reducing the incidence of primary breast malignancies, reducing secondary head and neck cancer, and enhancing regression of cervical IEN, these achievements have been modest, accompanied by considerable toxicity, and, most relevantly, been developed without the aid of prospective end point marker assessment. Despite being long time proponents of this approach, even we have begun to ask “why?” Put another way, “Does the Emperor have no clothes?”

Biomarkers have provided insight into the development of cancer and have been proposed to identify risk, target therapeutics, and identify responses to intervention. Whether biomarkers that identify putative risk for the development of cancer are likely to serve as SEBMs for the successful development of chemoprevention agents is the major focus and concern of this commentary.

Despite intense efforts to identify and verify candidate SEBMs, no prehistologic biochemical or molecular intermediate marker to date has been validated as a SEBM in animals or humans for any cancer (for review, see Ref. 4), and even the data with histological SEBMs (e.g., IEN) leave a lot to be desired. Schatzken and Gail (5) have recently reviewed the topic and express concerns similar to the current commentary. They begin their review by pointing out the savings in time and money validated SEBMs offer, only later discussing the fairly torturous journeys required to discover and validate the SEBMs. Thus the groundwork necessary to arrive at SEBM studies that are smaller and less expensive than observing frank clinical disease is daunting and may not be cost effective at the present state of knowledge of pathways to malignancy. Although one can make the argument that the problem is ineffective agents rather than unreliable markers, in fact many compounds are effective in animal models, and successes in reducing IEN and cancer in humans have been shown. The lack of real progress in cancer chemoprevention despite the enormous efforts expended compels us to carefully reconsider the theoretical constructs and primary assumptions that underlie strategies for understanding cancer evolution and chemopreventive agent development. Using head and neck cancer as a case example, a number of factors will be shown to significantly hinder our ability to validate SEBMs to develop candidate chemopreventive agents using SEBMs.

---

The Challenge of Validating SEBMs

The critical underpinning for SEBM development is an understanding of the biology behind the markers examined and their relationship to progression of the disease. We and others have discussed this subject comprehensively elsewhere (4), and a useful concise discussion of the whole issue of markers has been summarized recently (5). The utility of a SEBM has been posited as strongly related to its ability to predict subsequent development of cancer (6). If a SEBM (or group of SEBMs pooled together) does not accurately identify which individuals will develop cancer, there is little point in monitoring SEBM levels after drug intervention. For a SEBM to be useful to predict the effectiveness of a chemopreventive agent, the candidate compound must modulate it. Optimally, a SEBM should also be able to predict the likelihood of developing cancer, and, ultimately, modulation of biomarker levels by the compound should reduce the likelihood of developing cancer. To date, we know of few markers that are both necessary and sufficient prerequisites to cancer and are valid for entire, or even just sizable, populations. The best markers currently available are presence of dysplastic changes and development of IEN. The rate of malignant transformation of documented IEN for different tumor sites is variable, and for this reason, even IEN is not an ideal SEBM.

A substantial reason for the interest in SEBMs is that their use in chemopreventive agent development should make these studies smaller, shorter, and less expensive (e.g., see Ref. 5). This overall assumption is unlikely to be true. Clearly, studies could be shorter, given a validated SEBM, if shorter refers to the time spent with each subject in the study. However, the studies could only be smaller (fewer participants), compared with studies of direct clinical outcome if efforts are made to stack the deck up front, that is, find people with the critical

---

Received 1/6/03; revised 4/4/03; accepted 4/14/03.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 Supported in part by Grants U01-CA46496 and P30CA62203 from the National Cancer Institute.

2 To whom requests for reprints should be addressed, at Chao Family Comprehensive Cancer Center, 101 The City Drive South, Orange, CA 92868. E-mail: FLMeyoske@UCI.edu.

3 The abbreviations used are: SEBM, surrogate end point biomarker; IEN, intraepithelial neoplasia; LOH, loss of heterozygosity.
already validated SEBM and include only those people in the study. However, because sporadic cancers are rare, expression of a tightly coupled SEBM is likely to be rare, and much effort would go into screening (and rejecting) many prospective subjects. It is therefore not clear that a marker with high sensitivity and high specificity for malignancy could make studies smaller or even much less expensive if the screening process is counted in total study duration and cost. All this presupposes the marker has been validated, a large, lengthy undertaking in itself.

Alternatively, if the prospective SEBM is not tightly coupled to clinical outcome, then screening candidates for expression of the SEBM becomes progressively less worthwhile as the sensitivity and specificity of the marker decline, and the size of the study becomes driven by the clinical outcome in the wider population. The trial is not smaller any more, and when the budget is cut, it is the SEBM laboratory work and so forth that will have to go.

Multiple Pathways to Cancer Development
A major limitation to SEBM development and validation is that we cannot predict with a high level of accuracy who will develop cancer or identify markers that accurately predict an individual’s risk for developing cancer in the future. Without markers that can reliably discriminate which individuals will progress and which will not, the task of SEBM discovery and validation is at best tortuous. Accurately identifying which individuals are predisposed to develop sporadic cancer is a monumental challenge. Even in the presence of documented carcinoma in situ, not all lesions develop into invasive cancer (7). Why has identification of persons at risk for developing cancer based on measurement of SEBMs been so elusive? This strategy has been successfully used for cardiovascular disease, where cholesterol and triglycerides provide useful information not only for populations but also for a given individual at risk for atherosclerosis. No preclinical risk marker like cholesterol has been discovered for cancer, despite prolonged searching. Perhaps the failure to replicate the strategy successfully used for AIDS and cardiovascular drug development lies in part in the complexity of cancer development. The pathophysiology of atherosclerosis is relatively straightforward compared with that of carcinogenesis.

Multiple genetic injuries or “hits” are required to develop cancer. Recent studies of human cells in culture demonstrate, even under the most controlled circumstances, that targeted disruption of at least four distinct signaling pathways is required to produce malignant transformation in human cell lines (8). In adult humans, between 6 and 12 discrete hits are considered necessary to develop an invasive sporadic malignancy (9). If neoplastic progression results from multiple pathways, then a given marker of genetic damage to one pathway will not correlate with clinical outcome. Alternatively, if there were an alternative pathway to cancer that bypasses the candidate SEBM, then lack of expression of the biomarker would provide false reassurance. Thus any one marker will yield many false negatives and many false positives. This significantly limits the utility of any one marker for predicting cancer risk outside of a relatively small and hard-to-identify group of people or the ability of a single SEBM to predict effectiveness of a chemopreventive agent in sizable populations.

The mathematics of combinations may help illuminate the problem (10). If we assume biomarkers tightly coupled to pathways in a one-to-one fashion, then the number of distinct sets of biomarker patterns to be monitored increases dramatically with the number of pathways and the minimum number of required disruptions. As shown in Table 1 in the “Appendix,” if there are 10 relevant pathways, then there are 210 distinct patterns of biomarkers to monitor if disruption of any 4 pathways leads to malignancy. The number of distinct biomarker patterns possible can easily exceed the sample size of a Phase II trial. The task of validating each biomarker pattern for each agent under development could be daunting indeed.

Put another way, if six terrorists with different skills each take a separate road to meet to make a cancer “bomb,” and any two of them can assemble a plausible device, then there are 15 different devices potentially capable of producing malignancy (six choose two, see “Table 1”). Although our road maps are incomplete, we want to assign an “agent” to each road with the job of detecting and arresting the terrorist, infallibly. Partial success in this effort will not stop the production of bombs. If we knew which devices work and which, if any, are reliably duds, then perhaps we could focus our efforts to better effect. But we have much work to do to be able to tell the duds from the real dangers. We have to identify all of the terrorists and learn which combinations of their skills represent real danger before we can realize savings of effort, time, or treasure by focusing on those with real malignant potential and ignoring the rest. The actual biology is probably much more complex, and the combinatorial metaphor may be too simple, but the implications of the combinatorial metaphor are sobering nonetheless. It seems unlikely that one marker will emerge as a clinically useful surveillance tool for all populations and all times for most cancers. Given that no single marker is likely to be a perfect predictor of eventual malignant outcome, then statistical literature suggests the costs to “validate” a single marker may rival the costs of direct observation of the effects of agents on the clinical outcome (6). If validation requires that high correlations between surrogate and clinical outcome at the individual level be demonstrated across several populations, then the point of the exercise, to save time and effort, may well be lost.

Low Incidence of Adult Cancers Limits Predictive Value of Biomarkers
The difficulty of validating SEBMs is further illustrated by examination of problems associated with risk profiling of head and neck cancer. The baseline incidence of the disease is an important consideration when attempting to construct predictive models for disease progression. For common diseases such as coronary artery disease, well-described risk factors such as elevated cholesterol provide a useful and validated indicator of an individual’s risk for developing the disease, and modulation of the marker also correlates with change in risk (11, 12). Even though hypercholesterolemia is not required to develop coronary artery disease, and it can be argued that it is not tightly coupled to outcome at the individual level, cholesterol level is useful because the incidence of the disease in the population is very high, and decreasing cholesterol level decreases risk of mortality for many individuals and therefore in the population as a whole. Head and neck cancer and most other cancers differ in that although there are risk factors significantly associated with development of these malignancies in a population, an individual possessing the major risk factors still has a low personal likelihood of developing cancer because the baseline risk for the disease in the general population is very low. In comparison, the baseline risk for developing heart disease is very high.
Identification of Intermediate Markers in Premalignant Lesions

Extensive effort has been dedicated to identification of intermediate markers of risk for head and neck cancer. Because the likelihood of developing oral or pharyngeal cancer is much higher in patients with oral leukoplakia and erythroplakia (13–15), study of biomarkers in this group of patients has been undertaken in the hope of finding candidate SEBMs for further study in chemoprevention efforts. Several recent studies characterize the current state of intermediate marker identification and risk factor modeling of oral cancer. Rosin et al. (16) analyzed LOH at 19 microsatellite loci in archived oral leukoplakia biopsy specimens and correlated findings with tumor registry information. LOH at 3p and/or 9p plus an additional site predicted subsequent progression to invasive carcinoma of 53% (95% confidence interval, 38–74%). Lee et al. (17) reported 10-year study results of 71 patients with advanced oral premalignant lesions treated in chemoprevention trials of 13-cis-retinoic acid. They sought to identify predictive markers of malignant transformation that would supplement histological findings from biopsies of premalignant lesions (18). Twenty-one participants developed head and neck squamous cell carcinoma during the follow-up period. A combination of three markers (chromosomal polysomy, p53 expression, and LOH at 3p or 9p) produced a biomarker score that retrospectively predicted malignant transformation rates for patients with hyperplastic or mildly dysplastic lesions of 1 of 22 (4.5%), 6 of 22 (27.3%), 4 of 7 (57.1%), and 2 of 2 (100%) if the subject had respectively 0, 1, 2, or all 3 markers present (17). Sudbø et al. (19) prospectively demonstrated successful stratification of dysplastic oral epithelial lesions based on DNA ploidy determined by flow cytometry. The likelihood of developing oral cancer at or near the site of oral lesions ranged from 2.9% (3 of 105) for diploid lesions, 60% (12 of 20) for tetraploid lesions, and 78% (21 of 27) for aneuploid DNA.

These studies are representative of the best efforts thus far to identify markers associated with risk of developing oral cancer (for that matter, any cancer). The first two reports were retrospective analyses, which need prospective validation. The third was prospective analysis of a general marker of cellular disruption (DNA ploidy) in patients with confirmed dysplastic lesions. Unfortunately, none of the markers appears to have strong enough association with the development of cancer to have real potential of being validated as a SEBM, that is, be able to forecast with high accuracy the likelihood of developing disease. Inability to identify markers that can predict (or explain) a significant proportion of the likelihood of developing cancer does not bode well for SEBM development. Whether high-throughput measurement of multiple markers, such as that done in microarray expression evaluation, will improve the identification of useful markers remains to be seen, but the methodology and analysis are subject to the same constraints that currently impede discovery of successful SEBMs outlined herein. Perhaps, however, after the pathways to malignancy are better understood, and several (or many) markers or patterns of markers are identified (20), SEBMs will be useful in the design of improved chemoprevention agents (21).

IEN as a Target

Recently, it has been proposed that the prevention and treatment of IEN should serve as a generic target for accelerated new drug development (22). An accompanying commentary argued the case as cancer prevention by delay (23). Although IEN is considerably closer to malignancy than molecular or biochemical SEBMs, the natural history of IEN progression to malignancy in various organ sites is highly variable and, in most cases, of long duration. As proposed in these articles, the declaration is made that premalignancy (IEN) is a disease and is worthy of treatment in and of itself, which cuts through the entire SEBM argument. Because premalignancies have been treated by surgical means for decades, the observation that medical treatment would be worthwhile should come as no surprise. The development of IEN as an efficacious SEBM is nevertheless subject to the same limitations discussed herein for pre-IEN markers, although the constraints may be somewhat less severe. If indeed IEN is accepted as a target for medical intervention, then the pursuit of pre-IEN SEBMs as predictors of chemoprevention drug effect may not be worthwhile, and increased effort should be placed on identifying biochemical or molecular risk factors in IEN lesions that predict progression to malignancy because even histological preneoplasias are highly variable in their risk to develop clinical malignancy.

SEBM as Predictive, not Prognostic

It should be clearly understood that our concern with using SEBMs for chemoprevention agent selection is at the predictive rather than the prognostic level (SEBMs as measures of effectiveness of treatments rather than as heralds for cancer development). A prognostic marker can forecast the likelihood of developing cancer, whereas a predictive marker provides information about the effectiveness of a chemopreventive agent. It seems likely that prognostic factors will only be highly useful as predictors of malignancy when the former are integrally tied to the biology of the disease (e.g., estrogen receptor status in breast cancer; Ref. 24). How comprehensively the genetic paradigm evolved from studies of hereditary cancers will translate to sporadic cancer remains an open question, as discussed elsewhere (25). On the other hand, candidate agents that are tightly tied to the specific biochemistry of the intervention (e.g., polyamines and difluoromethylornithine; Refs. 26 and 27) may not affect the underlying biology sufficiently to be prognostic but may still be predictive. The distinction between prognostic and predictive is an important one that has been neglected or glossed over and has resulted in the measurement and over emphasis for SEBMs without the supporting context.

Conclusion

Although a number of risk factors have been identified, biomarkers have been measured, and the risks to populations identified by these markers have been well characterized, we do not appear to be close to having a validated SEBM for head and neck (or any other) cancer. The assumptions that we can identify molecular biomarkers on the several pathways to cancer to predict who will develop cancer and that SEBMs can be identified among these markers need critical reexamination.

If a (putative) marker may be compared with a population screening test, then any modern, introductory text in epidemiology will point out that utility is low when the prevalence of disease is low. The situation is made worse when the marker has low sensitivity (there are many false negatives) and low specificity (there are many false positives), as a consequence of alternate pathways and system redundancy. Not so elegant in theory as we once thought, in practice the concept has demonstrated very little success. As discussed in this commentary, a number of factors significantly hinder our ability to identify SEBMs for cancer that can be used to accurately measure the
Table 1  Number of ways to obtain a minimal number of disrupted pathways or loci on pathways (combinations) as a function of the number of pathways or loci subject to disruptions and the minimum number of disruptions required

<table>
<thead>
<tr>
<th>No. of pathways subject to disruption</th>
<th>Minimum no. of disrupted pathways yielding malignancy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>8</td>
<td>28</td>
</tr>
<tr>
<td>10</td>
<td>45</td>
</tr>
<tr>
<td>12</td>
<td>66</td>
</tr>
<tr>
<td>14</td>
<td>91</td>
</tr>
</tbody>
</table>

* Mathematically, the number of sets of r items that can be selected from a pool of N items, ignoring the order of selection (e.g., Ref. 10). nCr = n!/(r!(n – r)!).

effect of candidate chemopreventive agents and predict an individual’s risk of developing cancer. On the other hand, linking new agent development to the relevant biochemical perturbations is likely to predict whether a compound is worth pursuing. If an effect cannot be demonstrated in Phase II studies, then selection of such a compound for a definitive trial represents wishful thinking. Because of the limited power of Phase II studies, there is a significant possibility that a potentially effective agent will not be identified, which would argue for rather large and randomized (against placebo) Phase II trials in the drug development process. Meanwhile, someone hand the Emperor some sunscreen.

Acknowledgments
We thank Eugene Gerner, Ernie Hawk, and Jaye Viner for reading the manuscript and for their comments and suggestions.

Appendix
The mathematics of combinations is illustrated in Table 1.

References
Point: Surrogate End Point Biomarkers Are Likely To Be Limited in Their Usefulness in the Development of Cancer Chemoprevention Agents against Sporadic Cancers


**Updated version**
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/12/7/589

**Cited articles**
This article cites 22 articles, 6 of which you can access for free at:
http://cebp.aacrjournals.org/content/12/7/589.full#ref-list-1

**Citing articles**
This article has been cited by 1 HighWire-hosted articles. Access the articles at:
http://cebp.aacrjournals.org/content/12/7/589.full#related-urls

**E-mail alerts**
Sign up to receive free email-alerts related to this article or journal.

**Reprints and Subscriptions**
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

**Permissions**
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.