Review

Perspectives on Surrogate End Points in the Development of Drugs that Reduce the Risk of Cancer

Division of Cancer Prevention, National Cancer Institute, Bethesda, Maryland 20892 [G. J. K., K. M. J., C. W. B., P. G., J. A. C., E. T. H.], and CCS Associates, Mountain View, California, 94043 [C. C. S., L. A. D.]

Abstract

This paper proposes a scientific basis and possible strategy for applying surrogate end points in chemopreventive drug development. The potential surrogate end points for cancer incidence described are both phenotypic (at the tissue, cellular, and molecular levels) and genotypic biomarkers. To establish chemopreventive efficacy in randomized, placebo-controlled clinical trials, it is expected that in most cases it will be critical to ensure that virtually all of the biomarker lesions are prevented or that the lesions prevented are those with the potential to progress. This would require that both the phenotype and genotype of the target tissue in agent-treated subjects, especially in any new or remaining precancers, are equivalent to or show less progression than those of placebo-treated subjects. In the National Cancer Institute chemoprevention program, histological modulation of a precancer (intraepithelial neoplasia) has thus far been the primary phenotypic surrogate end point in chemoprevention trials. Additionally, we give high priority to biomarkers measuring specific and general genotypic changes correlating to the carcinogenesis progression model for the targeted cancer (e.g., progressive genomic instability as measured by loss of heterozygosity or amplification at a specific microsatellite loci). Other potential surrogate end points that may occur earlier in carcinogenesis are being analyzed in these precancers and in nearby normal appearing tissues. These biomarkers include proliferation and differentiation indices, specific gene and general chromosome damage, cell growth regulatory molecules, and biochemical activities (e.g., enzyme inhibition). Serum biomarkers also may be monitored (e.g., prostate-specific antigen) because of their accessibility. Potentially chemopreventive drug effects of the test agent also may be measured (e.g., tissue and serum estrogen levels in studies of steroid aromatase inhibitors). These initial studies are expected to expand the list of validated surrogate end points for future use. Continued discussion and research among the National Cancer Institute, the Food and Drug Administration, industry, and academia are needed to ensure that surrogate end point-based chemoprevention indications are feasible.

Introduction

Surrogate end point biomarkers are an important aspect of the chemopreventive drug development process (e.g., Refs. 1–7). This paper provides our present understanding of the value of these end points in chemopreventive drug development, along with a scientific basis and strategy for their present and future application. Cancer chemoprevention shares the interest and need for surrogate end points in drug development with other chronic diseases of aging (e.g., cardiovascular disease) and life-threatening diseases (e.g., AIDS). Particularly, the use of blood lipid lowering as a surrogate end point for cardiovascular disease provides a model for and insight into the issues that might surround the use of surrogates for cancer incidence in chemoprevention studies.

Cancer chemoprevention can be defined as treatment of carcinogenesis, i.e., its prevention, inhibition, or reversal (e.g., Ref. 5). In most epithelial tissues, accumulating mutations (i.e., genetic progression) and loss of cellular control functions are observed as the phenotype changes from normal histology to early IEN2 (2) to increasingly severe IEN, superficial cancers, and finally invasive disease. There are likely to be situations in which the process is relatively aggressive (e.g., in the presence of a DNA repair-deficient genotype or viral transformant such as the human papilloma virus), but generally, these changes appear to occur over a long time period (Fig. 1). For example, in the breast, it is estimated that progression from atypical hyperplasia through DCIS to adenocarcinoma may require 30 years or more (8, 9). Colorectal adenomas may form over a period as long as 5–20 years, and progression from adenoma to colorectal carcinoma may require another 5–15 years (10, 11). PIN may develop over ~20 years (12). The development from PIN to early latent cancer may take ≥10 years, and clinically significant carcinoma may not occur until 3–15 years later (12).

The prolonged time course of carcinogenesis provides an opportunity for chemoprevention—to intervene when the mutations are fewer, even before tissue level phenotypic changes are evident. However, the long latency also presents significant

1 The abbreviations used are: IEN, intraepithelial neoplasia; APC, adenomatous polyposis coli; CHD, coronary heart disease; CIN, cervical IEN; DCIS, ductal carcinoma in situ; EGFR, epidermal growth factor receptor; FAP, familial adenomatous polyposis; LOH, loss of heterozygosity; ODC, ornithine decarboxylase; PCNA, proliferating cell nuclear antigen; PIN, prostatic IEN; PSA, prostate-specific antigen; TGF, transforming growth factor; VPB, ventricular premature beat.

Received 6/3/99; revised 10/18/99; accepted 11/1/99.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 To whom requests for reprints should be addressed, at the National Cancer Institute, Division of Cancer Prevention, EPN 201, MSC 7322, 9000 Rockville Pike, Bethesda, MD 20892-7322

2 The abbreviations used are: IEN, intraepithelial neoplasia; APC, adenomatous polyposis coli; CHD, coronary heart disease; CIN, cervical IEN; DCIS, ductal carcinoma in situ; EGFR, epidermal growth factor receptor; FAP, familial adenomatous polyposis; LOH, loss of heterozygosity; ODC, ornithine decarboxylase; PCNA, proliferating cell nuclear antigen; PIN, prostatic IEN; PSA, prostate-specific antigen; TGF, transforming growth factor; VPB, ventricular premature beat.
challenges for the clinical phase of chemopreventive drug development. In this paper, we discuss a potential chemopreventive drug development strategy that addresses one major challenge—the long delays that could be expected, even in the presence of IEN, if modulation of a cancer end point were required to demonstrate efficacy as part of the usual process for drug approval (2–6). As suggested by the data presented above, cancers generally develop over decades, and IEN (e.g., PIN, colorectal adenomas) may also progress slowly (5–20 years). Typically, cancer incidence reduction trials have planned durations of 5–10 years, with anticipated accrual in the tens of thousands. We believe that the strategy we describe could allow demonstration of chemopreventive efficacy in most cancer targets in 3 years with several hundred, not tens of thousands, subjects and is possible because of increasing knowledge of the genetic, histopathological, and molecular basis of carcinogenesis. This strategy involves identifying, validating, and using phenotypic biomarkers (at the tissue, cellular, and molecular levels) and genotypic biomarkers as surrogate end points for cancer incidence.

Biomarkers of Carcinogenesis

IEN, such as colorectal adenomas, PIN, and CIN, are primary examples of tissue level phenotypic biomarkers that, because they are on the causal pathway to and are direct precursors of cancer, are generally considered for follow-up using molecular and genetic techniques (2, 3). Cellular biomarkers, such as nuclear and nucleolar morphology, mitotic index, and DNA ploidy, are also being evaluated in ongoing studies; they may be useful in characterizing the progression of IEN (2, 5). Other possibly useful genotypic biomarkers include LOH and gene amplification, either at specific gene loci (e.g., those for tumor suppressors such as p53 or tumor growth accelerators such as c-erbB2) or at panels of microsatellite loci where mutations indicate increasing genomic instability (13). Both phenotypic and genotypic changes during carcinogenesis may also be manifested by molecular biomarkers (3). For example, excess proliferation might be seen in increased levels of cellular antigens, such as PCNA or Ki-67/MIB-1 or overexpression of growth factors, such as epidermal growth factor, TGFα, and insulin-like growth factor I; reduced propensity to undergo apoptosis may be detected by increased expression of bcl-2. Aberrant differentiation may result in changes in G-actin, cytokeratins, and blood group antigens. Other molecular biomarkers may reflect general changes in cell growth control. These include TGFβ, cyclins, p53, and other tumor suppressors, as well as mutations and overexpression of oncogenes associated with carcinogenesis, such as ras and the transcription factors myc, fos, and jun. Tissue- and drug-related biomarkers may also be useful. Examples of tissue-related biomarkers are the expression of estrogen receptors in breast and PSA in prostate. Drug-related biomarkers associated with chemopreventive activity include inhibition of ODC by 2-difluoromethylornithine and inhibition of prostaglandin biosynthesis by nonsteroidal anti-inflammatory drugs.

Rationale for Using Biomarkers of Carcinogenesis as Surrogate End Points—Importance of Developing Molecular Progression Models

A key concept supporting the use of these biomarkers is that carcinogenesis is progressive. Progression has been mapped in target tissues by the appearance of specific molecular and more general genotypic damage associated with increasingly severe dysplastic phenotypes (e.g., Ref. 13 and other studies cited below). In many cases early, critical steps include inactivation of tumor suppressor genes, such as APC or the breast cancer (BRCA) gene and activation of oncogenes such as ras. Carcinogenesis may take multiple paths and be multifocal; not all cancers in a given tissue nor all cells in a given cancer may ultimately contain the same lesions. Progression may also be influenced by factors specific to the host tissue’s environment, such as the action of hormones produced in stroma around the developing epithelial tumor and changes in tissue structure (e.g., Refs. 14–16). Further, carcinogenesis may not necessarily
be driven by the order in which the changes appear; the disorganization caused by and exacerbating the accumulation of multiple effects may be more important. This disorganization is an obvious manifestation of carcinogenesis. Progression models that reflect increasing disorganization have been developed by Vogelstein, Sidransky, and their colleagues, the seminal work being that by Fearon and Vogelstein (17) in the colon. These researchers have also described carcinogenesis in the brain (18), bladder [Refs. 19–22; see also Simoneau and Jones (23)], and in the head and neck (13). Lam, Gazdar, and their colleagues (24, 25) have described early analyses of chromosomal loss correlating to grade of dysplasia and the appearance of non-small cell lung cancer. Also, Larson et al. (26) described accumulating chromosomal loss at defined loci in CIN. It is these genotypic and corresponding tissue and cellular histological lesions or biomarkers, when they are sufficiently stable to allow screening, are now available for adenomas ≤1 cm in highest potential to serve as surrogate end points. Specific carcinogenesis-associated molecular lesions identified thus far, although important, may not be the most informative among those that will be discovered as research continues. Thus, at present, because most cancer does not appear to occur unless it is preceded by an abnormal histological precancer phenotype, a focus on this abnormal phenotype, which integrates the relevant genetic and molecular changes, and, as described below, direct measurement of the accompanying genotypic changes, appears to us to provide the best opportunity for validating surrogate end points.

**Phenotypic and Genotypic Surrogate End Points to Establish Chemopreventive Efficacy**

We presently view IEN, the embodiment of the abnormal cancer phenotype, as a promising surrogate end point for clinical chemoprevention studies in epithelial tissues (2, 3, 27). Although shorter than the period for developing cancer, the latency for IEN progression can also be lengthy compared with the practical time frame for a chemopreventive intervention study. Importantly, the number of precancers may far exceed the number of cancers that subsequently develop in the target tissue, and behavioral (e.g., smoking history), environmental (e.g., hormonal status), and coexisting disease (e.g., immune system competence) factors may influence progression in individual subjects. IEN that will progress also may have particular characteristics predisposing them to develop into cancers. For example, the potential of colorectal adenomas to progress to cancer correlates to histological growth pattern, size, and severity of dysplasia (28–30). Two to five percent of tubular, 22% of tubulo-villus, and 20–55% of villous adenomas progress. Risk of malignancy is negligible for adenomas ≤1 cm in diameter and increases at larger diameters. Of the 70% of adenomatous polyps that are mildly dysplastic, no more than 5% progress to cancers; whereas, a much higher fraction (up to 55%) of the 10% of adenomas that are severely dysplastic become cancers. In one study, one-third of severely dysplastic adenomas contained invasive carcinoma (see Ref. 28), and severe dysplasia is found most frequently in larger adenomas with villous histology. Therefore, to establish chemopreventive efficacy, we expect that it will be critical to ensure that virtually all of the precancer lesions are prevented or that the lesions prevented are those with the potential to progress.

For these reasons, drug-induced prevention or regression of IEN determined histologically will not often be sufficient to determine chemopreventive efficacy. The evaluation should usually also consider the specific and general genotypic effects comprising the progression models for carcinogenesis and may also look at molecular pathology. As described in detail below, we believe determining that a reduced incidence of new precancers is chemoprevention requires that the genotype of the target tissue in agent-treated subjects, especially in any new precancers, is equivalent to or shows less genetic progression than that of placebo-treated subjects. Similarly, in studies with regression of existing precancers as the end point, where regression is incomplete, we feel that the remaining lesions in the agent-treated subjects should have genotypes equivalent to or showing less progression than placebo control subjects.

A critical issue to the application and validation of these surrogate end points is developing standardized, appropriate, and quantitative techniques for sampling the target tissues. Improved diagnostic tools such as gene-chip analyses, the confocal microscope, digital mammography, the lung-imaging fluorescence endoscope for visualizing bronchial tissue, and the magnifying endoscope for colorectal monitoring will be critical to assuring the adequate visualization and monitoring of precancerous tissue.

**Cohorts for Surrogate Endpoint Chemoprevention Studies**

Another important concept is the definition of high-risk tissue, particularly as applied to patients with previous cancers or precancers. Generally, these patients show an increased risk for developing new primary lesions in tissue that is histologically related to tissue from which the original lesion arose. Slaughter (31) coined the term “field cancerization” to describe the early evidence of carcinogenesis found in normal appearing mucosa of patients with previous head and neck cancers. In fact, the lifetime risk for a second primary tumor of the aerodigestive tract following a squamous cell cancer of the head or neck has been estimated at 20–40% (32). Many studies, particularly those carried out by Hong, Lippman, Hittelman, and their colleagues (32–34), have confirmed this phenomenon. For example, Hittelman (33) delineated the use of chromosome in situ hybridization to detect carcinogenesis-associated genotypic changes (≥3 copies of a single chromosome) in normal and precancerous tissue nearby head and neck cancers. The tissues were histologically and otherwise phenotypically distinct from cancers; hence, the genotypic changes did not likely result from random sampling errors. In these studies, the degree of genetic change detected correlated to histological progression of the lesion toward cancer. Very importantly, 8 of 15 patients (~53%) having premalignant lesions of the oral cavity containing high levels of genetic damage (~3.5% of cells with three or more copies of chromosome 9) subsequently developed aerodigestive tract cancer compared with none among patients with lower levels. Similar results were found by Hittelman and his colleagues (33) at chromosome 9 in lung tissue from previous smokers and at chromosome 17 from breast tissue (35), and by Segers et al. (36) at chromosome 1 from cervical tissue from patients with various grades of CIN. Although none of these studies tracked the development of specific lesions into cancers, they all confirmed that carcinogenesis could be detected by genotypic changes in high-risk tissue.

The implication for clinical chemoprevention studies is that patients with previous cancers or precancers provide cohorts who are at high risk for new primary cancers and will benefit from chemoprevention, allowing smaller trials to be designed in which premalignant changes (surrogate end points) can be followed both phenotypically and genotypically. One criterion we use for selecting these cohorts is expectation of a
high incidence of the cancer or precancer, or observable progression of the precancer, under study within a reasonable time period. For Phase II and III studies using surrogate end points, regression of the precancer, under study within a reasonable time high incidence of the cancer or precancer, or observable progression may also be used to define high-risk patients. Studies are described in the text. Results are change (Δ) in end points from baseline to posttreatment; treatment groups are compared with placebo groups.

### Study/Case

**Prevention of Colorectal Adenomas**

<table>
<thead>
<tr>
<th>Case</th>
<th>Phenotype (primary end point)</th>
<th>Genotype</th>
<th>Evidence for chemopreventive efficacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50% fewer new adenomas in treatment group than placebo controls. Reduced incidence is statistically significant (P &lt; 0.05).</td>
<td>% LOH at chromosomes 5q, 17p, and 18q and ras mutations in tissue samples from treatment group and placebo controls not significantly different (P &lt; 0.05), or pattern of LOH is consistent with less progression in treatment group.</td>
<td>Treatment is chemopreventive. Genotype analysis is confirmatory evidence that prevention of new adenomas is general, and not selective inhibition of those less likely to progress.</td>
</tr>
<tr>
<td>2</td>
<td>50% fewer new adenomas in treatment group than placebo controls. Reduced incidence is statistically significant (P &lt; 0.05).</td>
<td>% LOH at chromosomes 5q, 17p, and 18q and ras mutations in tissue samples from treatment group significantly higher (P &lt; 0.05) than in placebo controls, or pattern of LOH is consistent with greater progression in treatment group.</td>
<td>Treatment is not conclusively chemopreventive. Genotype analysis suggests that treatment is inhibiting adenomas less likely to progress.</td>
</tr>
</tbody>
</table>

**Regression of Oral Dysplasia**

<table>
<thead>
<tr>
<th>Case</th>
<th>Phenotype (primary end point)</th>
<th>Genotype</th>
<th>Evidence for chemopreventive efficacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>50% more complete or partial (e.g., 50% reduction in size) regression of dysplastic lesion.</td>
<td>% LOH at chromosomes 9p21, 3p21, 17p13, 11q13, 13q21, and 14q31-32.1 in tissue samples from treatment group and placebo controls not significantly different (P &gt; 0.05), or pattern of LOH is consistent with less progression in treatment group.</td>
<td>Treatment is chemopreventive. Genotype analysis is confirmatory evidence that the regression observed is general and not just selective for dysplasia not likely to progress.</td>
</tr>
<tr>
<td>4</td>
<td>50% more complete or partial (e.g., 50% reduction in size) regression of dysplastic lesion.</td>
<td>% LOH at chromosomes 9p21, 3p21, 17p13, 11q13, 13q21, and 14q31-32.1 in tissue samples from treatment group significantly higher (P &lt; 0.05) than in placebo controls, or pattern of LOH is consistent with greater progression in treatment group.</td>
<td>Treatment is not conclusively chemopreventive. Genotype analysis suggests that treatment is regressing lesions less likely to progress.</td>
</tr>
</tbody>
</table>

### Modulation of Surrogate End Points that Provides Evidence for Chemoprevention

Important issues in using surrogate end points in chemoprevention studies are determining when a less-than-complete phenotypic response of a precancerous lesion constitutes prevention and the role of genotypic biomarkers in evaluating this result. A related, subtle efficacy issue is distinguishing chemoprevention from regression of existing disease. We believe that establishing chemopreventive efficacy solely on the basis of phenotypic regression would in most cases require near complete regression of existing lesions. Similarly, phenotypic chemoprevention would in most cases only be demonstrated rigorously by near complete inhibition of new lesions. With less than this level of efficacy, it is possible that the remaining lesions are those that will continue to progress to cancer.

However, if a posttreatment lesional genotype showed decreased incidence of cancer-related changes (either in specific genes or in more general measures of genomic instability) compared with the baseline, we feel that a significantly less than near complete regression could be considered prevention. Also, less than complete inhibition of phenotypic progression could establish chemopreventive efficacy if no lesions in the active intervention group exhibited a genotype that had progressed beyond those of the baseline lesions or placebo controls. This outcome would be further supported if the genotype of normal appearing tissue in the target also was stable or showed reduced cancer-related changes compared with the baseline. These concepts are summarized by the four hypothetical examples in Table 1. Each represents the results of using a potential chemopreventive agent to treat patients with precancer or with previous precancer or cancer. It is assumed that the phenotypic inhibition or regression observed is statistically significant and that previous studies have determined important cancer-related genotypic changes for this target.

Cases 1 and 2 are based on prevention of colorectal ade-
nomas as the primary end point. In both examples, patients receive baseline colonoscopies during which all visible adenomas are counted, measured, and removed. Biopsy samples are taken from polyps and normal appearing tissue for biomarker analysis. Patients are then randomized to treatment with a chemopreventive agent or placebo. After treatment, the patients again receive colonoscopies. Any new polyps are counted, measured, and excised, and, as at the baseline, biopsy samples are taken from both the polyps and normal appearing mucosa. In all patients, samples are taken from the same area of the colorectum at the baseline and posttreatment, and the biopsy sampling pattern is representative of the total tissue and is the same at both time points. In both cases, significantly fewer polyps (e.g., 50% lower incidence) are found in the treatment group than in the placebo group. However, the genotypic status of colorectal polyps and tissue in the treatment groups compared with the baseline differs. The ten loci that differ the most between the two cases, LOH at markers on chromosomes 5q [affecting the APC and the mutated in colorectal carcinogenesis (MCC) genes], 17p (p53 gene), and 18q [deleted in colorectal carcinoma (DCC) gene], along with ras mutations, have been shown to increase with progression of colorectal dysplasia (43). In case 1, the percentages of LOH at 5q, 17p, and 18q and of ras mutations are comparable in the treatment and control groups; also, posttreatment percentages are comparable to the baseline. Because the relevant genotypic biomarkers did not show progression, we feel that this study would provide evidence of chemopreventive efficacy. In case 2, the percentages of allelic loss at 5q, 17p, and 18q are significantly higher in polyps from the treatment group than from the placebo group. These data suggest that the treatment may only have inhibited less severe dysplasia, which would be less likely to develop into carcinoma. We would not consider this result a convincing demonstration of chemopreventive efficacy.

Cases 3 and 4 are based on the genetic progression model for head and neck cancer described by Sidransky and colleagues (13), showing the correlation of LOH frequencies at ten specific microsatellite loci and increasing numbers of affected loci to severity of dysplasia. In both studies, a chemopreventive agent or placebo is administered to patients with dysplastic lesions in the oral cavity. These lesions are measured at the baseline and are biopsied along with normal appearing adjacent tissue. Biopsies of both the dysplastic lesions and normal appearing tissue are also taken at the end of treatment. The same rigorous attention to adequate and representative sampling described for the colorectal adenoma studies is applied to these assessments. At the end of both studies, significant regression of the dysplastic lesions is observed. Fewer lesions remain in the treatment group than in the control group; lesions in the treated group are also significantly smaller. LOH of the ten important loci is analyzed in cells from the lesions and nearby, normal appearing tissue. In case 3, LOH frequency distribution is similar in the dysplasia from the treated and control groups. Also, LOH frequency distributions in normal appearing mucosa are comparable in all groups. Hence, assuming that the ten loci analyzed are informative biomarkers, no genetic changes are seen that would suggest that the lesions in the treated patients are likely to progress faster than those in controls. We expect that such a result provides supporting evidence that the reduced incidence of dysplasia is a true chemopreventive effect. Case 4 is the same hypothetical study with negative results. In this case the phenotypic regression is misleading. Although the number of lesions is fewer, the average frequency of LOH is significantly higher in lesions from the treated patients than the controls, indicating that the lesions in the treated patients are likely to progress more quickly than those in the controls and that the chemopreventive intervention is only preventing lesions less likely to progress. It should be noted that the reliability of such genotypic assessments would be determined by knowledge of the important genetic lesions.

Remarkable advances in genome sequencing and functional genomics and proteomics are being made that will soon produce genetic progression models that are increasingly more comprehensive and informative; these capabilities have been reviewed by Brown and Botstein (44). Besides the research to develop genetic progression models described above, the sequencing and functional analysis efforts of the Cancer Genome Anatomy Project are a major contribution to this knowledge. We envision two types of genotypic analyses that, after extensive sampling, methods, and statistical standardization, could be used as end points for chemoprevention studies. The first is monitoring prespecified sets of genetic lesions that are strongly associated with neoplastic progression and is analogous to and an extension of the LOH analysis cited in the hypothetical colorectal and oral cavity examples. Many academic researchers and commercial sources are now designing and producing gene chips (e.g., cDNA microarrays) that can be used to measure specific cancer-related genotypic changes. The second type of method is also based on microarray analysis, but is a more generalized comparison of gene expression in posttreatment and baseline lesions. One such method, cluster analysis of genome-wide expression patterns has been described by Eisen et al. (45).

Precedents for Surrogate Endpoints in Development of Drugs for Disease Prevention—Lipid Lowering Drugs in Prevention of Cardiovascular Disease

To date, the best characterized surrogate end points in drug development have been for AIDS (46, 47) and cardiovascular drugs (reviewed in Refs. 48 and 49). Because of the long time required for disease development, the multiple paths by which the disease progresses, and the chronic administration of preventive drugs, the course of cardiovascular disease closely parallels carcinogenesis. In the cardiovascular setting, a prominent surrogate end point is cholesterol level, which is a validated predictor of CHD (50). Modulation of cholesterol levels has been used to gain marketing approval for 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitors such as lovastatin (51, 52), simvastatin, pravastatin (53, 54), and gemfibrozil (55). 3-Hydroxy-3-methylglutaryl coenzyme A reductase catalyzes a critical step in cholesterol biosynthesis, the formation of mevalonate. Gould et al. (56) have carried out a meta-analysis of 35 randomized clinical trials that essentially summarizes the evidence supporting cholesterol lowering as a surrogate end point for CHD. The trials reviewed (all primary or secondary intervention studies of >2 years’ duration) include single drug studies such as the Helsinki Heart Study of gemfibrozil (55), as well as diet (57, 58), surgical (59), and multifactorial interventions (60, 61). The results show that cholesterol lowering is correlated to CHD, non-CHD, and overall mortality. Specifically, it was found that for every 10% lowering of cholesterol, CHD mortality was reduced by 13% (P < 0.002) and total mortality by 10% (P < 0.03), whereas no effect was found on non-CHD mortality. A caveat applies here as to all studies with biomarkers—the relationship between lower CHD and lower cholesterol is the result of averaging individual responses. There are many individuals for whom the correlation is not seen. In other words, when the parameter evaluated is one of several in the multifactorial disease process, other variables...
Review: Surrogate End Points in Chemopreventive Drug Development

Table 2: Validation of surrogate end points (relationship of surrogate end point to disease, effect of drug intervention on surrogate end point, effect of drug intervention on disease): cholesterol lowering and CHD

<table>
<thead>
<tr>
<th>Serum Cholesterol</th>
<th>Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Quartile (≥ 193 mg/100 ml)</td>
<td>57</td>
</tr>
<tr>
<td>2nd Quartile</td>
<td>85</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>102</td>
</tr>
<tr>
<td>4th Quartile (≥ 250 mg/100 ml)</td>
<td>152</td>
</tr>
</tbody>
</table>

Effect of drug intervention on cholesterol level, leading to approval by FDA for prevention of CHD (cited in Ref. 56). Simvastatin decreased serum cholesterol averaging 26% compared with placebo controls.

Effect of cholesterol lowering intervention on CHD, meta-analysis of trials ≥ 2 yr in duration (56).
For every 10% cholesterol lowering, CHD mortality reduced by 13% (P < 0.002) and total mortality reduced by 10% (P < 0.03).

along with secondary or indirect processes contributing to the disease (in the case of heart disease, examples are smoking history and diabetes mellitus) or protecting the subject may confound interpretation. Also, the proportion of disease attributable to a specific biomarker may vary according to this variety of intervening variables characteristic of any particular population.

The research on cholesterol lowering and other surrogate end points for cardiovascular disease provides a model for studies needed to validate surrogate end points—both in the relationship of the surrogate end point to the ultimate disease and in its ability to predict activity of a given drug on that disease. Table 2 shows data that validates cholesterol lowering as a surrogate end point for CHD; analogous data might be applied to a surrogate for cancer incidence. For example, it has been well-established that the presence of colorectal adenomas increases the risk for colorectal cancer (29, 40). It has been estimated that 2–5% of all colorectal adenomas progress to adenocarcinomas if not removed or treated (the rate increases with size and severity of dysplasia; up to 22% of tubule-villus colorectal adenomas progress; Refs. 10, 11, and 29). These data alone or combined with results of studies showing cancer risk reduction on polyp removal could possibly be used to validate adenomas as a surrogate end point for cancer incidence and demonstrate that a drug that prevents polyps has cancer chemopreventive efficacy. As will be discussed below, the cardiovascular and AIDS drugs also demonstrate problems associated with drug evaluations based on surrogate end points.

Issues in Using Surrogate End Points

There are several philosophical and practical issues that arise in applying surrogate end points to the evaluation of drug efficacy and, specifically, chemopreventive efficacy. Temple (49, 62) previously addressed many of these issues in the context of cardiovascular drug development.

Impact of Surrogate End Point Modulation on Cancer Incidence/Mortality. One issue is the degree of clinical benefit that should be derived from efficacy against the surrogate end point (48, 63, 64). As described by Blue and Colburn (63), surrogate end points fall onto a continuum from showing no particular clinical benefit but only correlation to the target disease end point (e.g., drug effect markers), through demonstrating a clinical benefit that is not a direct effect on the target disease (e.g., immunostimulation), to demonstrating a clinical benefit that is directly related to the target disease (e.g., inhibiting colorectal adenomas). Initially, the criteria we describe for selecting surrogate end points support drugs with a clinical benefit directly related to cancer incidence prevention. However, as more data are developed on the role of general genotypic and specific molecular changes in carcinogenesis and with careful correlational studies, the effects on surrogate end points with antecedent impact on clinical outcome may also support chemopreventive drug efficacy.

Quality of Life. Chemopreventive drugs may ultimately be given to asymptomatic populations for years or decades. Therefore, minimal toxicity is essential. Determining standards in terms of allowable type and frequency of side effects and impact on quality of life will be critical issues as chemopreventive drugs are introduced.

Adverse Effects May Not Be Observed in Short-Term Surrogate End Point Studies. It is also possible that life-threatening toxicities compromising the long-term use of drug would not be detected within the time frame of surrogate end point-based efficacy trials. In the meta-analysis of cholesterol-lowering interventions cited above, the investigators found that despite their cholesterol-lowering efficacy, fibrates such as gemfibrozil were associated with increased non-CHD mortality by ~30% (P < 0.01) and total mortality by ~17% (P < 0.01) on long-term administration (56). A different but dramatic example of unanticipated late toxicity is provided by the results of the Cardiac Arrhythmia Suppression Trial (48). This randomized, placebo-controlled trial of three type 1C antiarrhythmics was designed to evaluate mortality reduction in patients experiencing 10 VPBs/h and few or no symptoms following a recent myocardial infarction. Entry in the trial required that the patients respond to antiarrhythmic therapy as measured by at least a 70% reduction in VPBs as a surrogate for arrhythmia. This trial was stopped when it was found that drug treatment was associated with increased mortality or cardiac arrest despite lowering VPBs (65, 66).

Potential Surrogate End Points at Major Cancer Target Organs

We have previously described our criteria for selecting surrogate end points for clinical chemoprevention studies (see Table 3 and Ref. 3). There are now >40 clinical studies sponsored by our National Cancer Institute, Division of Cancer Prevention chemoprevention program, which are in progress and involve the evaluation of potential surrogate end points. These are primarily Phase II trials in 10 major cancer target sites (prostate, breast, colon, lung, head and neck, bladder, cervix, esophagus, skin, and liver). Table 4 surveys the cohorts and surrogate end points presently under evaluation in these studies. Usually, the primary end point is a histological modulation of a precancer. This modulation may be evaluated by both classical pathological techniques and by morphometry and cytophotometry using computer-assisted image analysis.

Additionally, we give high priority to biomarkers measuring specific and general genotypic changes correlating to the carcinogenesis progression model for the targeted cancer. Progressive genomic instability as measured by LOH or amplification at specific microsatellite loci was used by Sidransky and colleagues (Ref. 13; see also references to other aspects of this study elsewhere in this paper) to characterize head and neck carcinogenesis. These biomarkers are potential surrogate end points in head and neck and may also prove useful in other tissues where microsatellite instability is a predominant feature.
of carcinogenesis—for example, in hereditary non-polyposis colorectal cancer-associated and some sporadic colorectal cancers (67, 68). For all of the biomarkers, we feel that it is highly desirable to measure modulation quantitatively as the difference (Δ) between the biomarker value at the end of the treatment and the baseline. The change in the surrogate end point measures on chemopreventive treatment should also be compared with that seen in appropriate controls. Thus, we also believe that baseline biopsies or other tissue measurements are essential.

Although the precancer histological phenotype with the accompanying genotypic changes determined to be relevant in the genetic progression models serves as our primary focus for initially proving chemopreventive efficacy, other potential surrogate end points that may occur earlier in carcinogenesis are being analyzed in these precancers and in nearby normal appearing tissues. These biomarkers include proliferation and differentiation indices, specific gene and general chromosome damage, cell growth regulatory molecules, and biochemical activities (e.g., enzyme inhibition). Although these biomarker studies focus on characterizing effects in the cancer target tissue, serum biomarkers may also be monitored (e.g., PSA) because of their accessibility. In some cases, biomarkers specifically related to the postulated chemopreventive drug effect of the test agent are measured. Although such biomarkers do not necessarily demonstrate a chemopreventive effect, they are useful in determining that a biologically active dose of the agent was present and in evaluating the chemopreventive mechanisms that are operating. For example, in studies with the ODC inhibitor 2-difluoromethylornithine, tissue and serum polyamine levels are determined, and in studies with steroid aromatase inhibitors, tissue and serum estrogen levels are analyzed. These initial studies with drug effect biomarkers may expand the list of validated surrogate end points for future use.

A potential set of criteria for establishing chemoprevention in terms of the histological phenotypic and genotypic changes observed was described above. In Table 5, this description is extended to define the specific phenotypic and genotypic results from chemoprevention trials in the major cancer targets that we would consider definitive evidence of efficacy. Those results, which would not likely be considered sufficient to be definitive, but would provide significant supporting evidence, are also described. In the future, it is likely that many biomarkers that are now being investigated will be demonstrated to be predictive of cancer incidence in clinical, or in the nearer term, in animal studies. The potential role of such data are also cited in Table 5. Note that the focus in these tables is optimal results. In some cases, the trial designs adequate to provide such results have not yet been fully defined.

### Clinical Benefit Based on Surrogate End Points (Not Cancer Prevention)

There are several situations in which the treatment of precancerous lesions would appear to provide clinical benefit, notwithstanding the potential for cancer prevention. These benefits include reduced morbidity, enhanced quality of life, delayed surgery, and increased intervals for surveillance requiring invasive procedures.

#### Prevention of Precancers in Subjects at High Risk Associated with Genetic Predisposition (e.g., Prevention of Colorectal Adenomas in Patients with FAP)

FAP is characterized by germ-line mutations in the APC tumor suppressor gene. Usually starting when they are teenagers, patients with FAP develop hundreds of colorectal adenomatous polyps. If untreated, FAP patients will almost certainly develop colorectal cancer by age 50 (69); they are also at risk for developing other lesions, particularly duodenal polyps and cancers. Once adenomas begin to appear, these patients are monitored by periodic colonoscopy (at ~6-month intervals), removal of existing polyps, and cancer screening. When the polyp burden becomes unmanageable, most patients have partial or total colectomies. Thereafter, they continue to be monitored. Agents that prevent or slow the progression of the adenomas could benefit these patients by delaying the need for colectomy and increasing the intervals between surveillance colonoscopies and cancer screenings.

#### Prevention of Precancers for Which Organ Removal or Other Major Surgery with High Morbidity Is Standard of Care (e.g., Barrett’s Esophagus, Superficial Bladder Cancers)

Present treatment for Barrett’s esophagus, a precursor of esophageal cancer, may involve partial or total esophagectomy (70). Because of the high rate of their recurrence and potential for progression, treatment for superficial bladder cancers includes periodic surveillance (every 3 months) and removal of new lesions and may include cystectomy (71). In both diseases, treatment has profound detrimental effects on quality of life. Both are examples of situations in which preventive agents could provide a clinical benefit by reducing the need for these surgeries.

#### Prevention of Precancers in Patients at Risk for Recurrence (e.g., Sporadic Colorectal Adenomas)

New adenomas occur within 1–3 years after resection in ~30% of patients with sporadic colorectal adenomas or cancers (30). These patients...
Table 4  Potential surrogate end points at major cancer target sites being evaluated in Phase II/III studies sponsored or funded by the National Cancer Institute, Division of Cancer Prevention

<table>
<thead>
<tr>
<th>Target/cohort(s)</th>
<th>Primary end point</th>
<th>Histological/genotypic end point(s)</th>
<th>Other potential end points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prostate Patients with PIN</td>
<td>PIN regression/prevention</td>
<td>Histopathology: nuclear morphometry (nuclear texture and shape, [roundness]); nuclear morphology (size and number of nuclei); DNA ploidy</td>
<td>Proliferation (MIB-1 and proliferating cell nuclear antigen)</td>
</tr>
<tr>
<td>Presurgical early prostate cancer patients</td>
<td>Genotype: chromosome 8p LOH and 8q amplification (85), Fourier transform-infrared spectroscopy of DNA structure (86)</td>
<td>Apoptosis (number of apoptotic bodies, transglutaminase, bcl-2)</td>
<td></td>
</tr>
<tr>
<td>Breast Atypical hyperplasia, previously treated LCIS, DCIS or minimally invasive breast cancers</td>
<td>Hyperplasia/DCIS prevention/regression</td>
<td>Histopathology: mammographic density (87)</td>
<td>Proliferation (MIB-1 PCNA)</td>
</tr>
<tr>
<td>Atypical hyperplasia/epithelial hyperplasia and one or more biomarker abnormalities</td>
<td>Nuclear morphometry, DNA ploidy</td>
<td>Apoptosis (bcl-2)</td>
<td></td>
</tr>
<tr>
<td>Presurgical mammographically detected DCIS and minimally invasive cancers</td>
<td>Cell regulatory molecules (EGFR/c-erbB-2, estrogen receptor, IGF-1 (88, 89), p53)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colon FAP Prevention of colorectal adenomas</td>
<td>Histopathology: nuclear morphometry (nuclear texture and shape), nuclear morphology (size and number of nuclei), DNA ploidy</td>
<td>Proliferation (expansion of the proliferative compartment in colon crypts measured, for example, by BrdU uptake, S-phase fraction, PCNA, MIB-1, ratio of proliferation to apoptosis)</td>
<td></td>
</tr>
<tr>
<td>HNPPC Previous colorectal cancers, previous/current colorectal adenomas</td>
<td>Genotype: chromosome 5, 17, 18 LOH and ras mutations (43), microsatellite instability, DNA methylation pattern</td>
<td>Apoptosis (apoptotic bodies by confocal laser microscopy, TUNEL assay)</td>
<td></td>
</tr>
<tr>
<td>Lung Chronic smokers with proven bronchial dysplasia</td>
<td>Bronchial dysplasia regression</td>
<td>Histopathology: nuclear morphometry (pleomorphism), DNA ploidy</td>
<td>Proliferation (PCNA)</td>
</tr>
<tr>
<td></td>
<td>Genotype: chromosome LOH at LOH at 3p21, 3p24–25, 5q and 9p (e.g., Refs. 24 and 25) and in FRA3B/FHIT gene (Refs. 24, 25 and 91), mutagen sensitivity</td>
<td>Apoptosis (bcl-2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cell regulatory molecules (telomerase, EGFR, p53)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bladder Ta, T1 superficial bladder cancer</td>
<td>Prevention of new tumors</td>
<td>Histopathology: DNA ploidy</td>
<td>Proliferation (PCNA, M344)</td>
</tr>
<tr>
<td>Ta, T1 ≤ TIS superficial bladder cancer treated with BCG</td>
<td>Genotype: chromosome LOH in urine (Refs. 22 and 93)</td>
<td>Differentiation [G-actin (Refs. 92, and 93)]</td>
<td></td>
</tr>
<tr>
<td>Head and neck Previous head and neck cancers</td>
<td>Prevention/regression of dysplastic lesions</td>
<td>Histopathology: DNA ploidy</td>
<td>Proliferation (PCNA, MIB-1)</td>
</tr>
<tr>
<td></td>
<td>Genotype: Chromosome LOH (Refs. 13, 24, and 25)</td>
<td>Cell regulatory molecules (EGFR, c-erbB-2, TGFα, TGFβ)</td>
<td></td>
</tr>
<tr>
<td>Cervix CIN II/III</td>
<td>CIN regression</td>
<td>Histopathology: nuclear morphometry (pleomorphism, DNA content), DNA ploidy</td>
<td>Proliferation [PCNA, (Ref. 94)]</td>
</tr>
<tr>
<td></td>
<td>Genotype: chromosomal LOH (26)</td>
<td>Differentiation (keratins)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cell regulatory molecules (EGFR, ras expression/mutation)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Esophagus Barrett’s esophagus</td>
<td>Prevention/regression of Barrett’s dysplasia</td>
<td>Histopathology: nuclear morphometry (pleomorphism, DNA content), nuclear morphology (size and number of nuclei), DNA ploidy</td>
<td>Proliferation [PCNA, MIB-1)]</td>
</tr>
<tr>
<td></td>
<td>Apoptosis</td>
<td>Cell regulatory molecules (EGFR, p53)</td>
<td></td>
</tr>
<tr>
<td>Skin Actinic keratosis</td>
<td>Prevention/regression of actinic keratosis</td>
<td>Genotype: p53 (95)</td>
<td>Proliferation (PCNA, ODC activity)</td>
</tr>
<tr>
<td>Previous nonmelanoma skin cancer</td>
<td>Cell regulatory molecules (EGFR, TGFβ, p53)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* See references 2, 3, and 5 for general discussions on these studies. See also specific references cited in the table.

© 2000 American Association for Cancer Research.
Definitive
In a randomized, double-blinded, placebo-controlled trial, where the target tissue has been adequately sampled:
- Near complete prevention of IEN in the treatment group(s) that is statistically significant compared with the control group
- Statistically significant reduction in incidence of IEN in treatment vs. control group and evidence from histopathologic and/or genotypic analysis that IEN in both groups are equivalent in progression or that those in treatment group are less severe than those in the control group
- Near complete regression of existing IEN in treatment group that is statistically significantly greater than that in the control group
- Statistically significant increase in IEN regression in treatment group compared with control group and evidence from histopathologic and/or genotypic analysis that remaining IEN in both groups are equivalent in progression or that those in the treatment group are less severe than those in the control group

Supporting
In a randomized, double-blinded, placebo-controlled trial, where the target tissue has been adequately sampled:
- Statistically significant reduced incidence or increased regression of IEN in the treatment group compared with control group, without confirmatory evidence from histopathologic and/or genotypic analysis that IEN lesions remaining in both groups are equivalent in progression or that those in treatment group are less severe than those in the control group
- Statistically significant modulation of one or more histologic, genotypic, or molecular end points in the treatment group compared with the control group, where such modulation is strongly correlated to reduction in cancer or IEN incidence. This modulation should be observed in a well-characterized, standardized, quantitative assay. Its validation may be obtained in an animal model of the target cancer

Future Definitive
In a randomized, double-blinded, placebo-controlled trial
- Statistically significant modulation of one or more histologic, genotypic, or molecular end points in the treatment group compared with the control group, where such modulation is strongly correlated to reduction in cancer incidence and/or IEN incidence and grade. This modulation should be observed in a well-characterized, standardized, quantitative assay. This end point has been validated as a predictor of reduced cancer incidence or IEN incidence and grade in clinical studies

This table delineates science-based criteria for evaluating chemopreventive efficacy based on surrogate end points. Practical application of these criteria in a drug development setting will be most successful where the genetic progression model and the technologies to analyze the genetic and histopathologic changes are well developed. Also, other factors will undoubtedly influence the weight such evidence contributes to decisions on marketing approval. Particularly, the intended indication and other human experience with the drug—e.g., safety, epidemiological, or other clinical evidence of efficacy in the target population—could be significant.

Potential Application of Surrogate Endpoints in Gaining Marketing Approval for Chemopreventive Drugs
This paper focuses on what we believe are the critical scientific aspects of developing surrogate end points to characterize cancer chemopreventive efficacy. We anticipate that the material presented will serve as a basis for designing clinical development strategies to gain marketing approval for chemopreventive drugs. As we’ve discussed here and previously (1, 2, 5, 6), the multipath, multifocal nature of carcinogenesis, as well as the very small percentage of early lesions that progress to cancers and the long time required for cancers to develop, suggests that, initially, the most successful strategies will use well-defined precancers (IEN) as the surrogate end points for cancer incidence. Despite their close temporal and histological association with cancers, only a relatively small percentage of IEN will progress. Therefore, the determination of chemopreventive efficacy will rely on the assurance that the lesions most likely to progress are inhibited (e.g., the genotype of any posttreatment lesions is equivalent to or indicative of less progression than baseline lesions). The phenotypic changes seen in IEN during short-term studies are likely to be subtle; therefore, quantitative measurements such as computer-assisted image analysis are desirable; similarly, the evaluation of genotypic changes requires sensitive, quantitative analysis of gene expression such as that afforded by the various DNA microarray techniques. We also recognize that standardization is critical, including determining adequate sampling, handling of nonrelated biopsy effects, and timing of biomarker assessment relative to normal biological cycles (e.g., timing for measurement of breast cell proliferation during the menstrual cycle). The gold standard for validating surrogate end points is comparison with cancer incidence reduction. The resources (e.g., time and number of subjects) required to successfully complete such validation are enormous. We believe that continued discussion and research on alternative strategies among the National Cancer Institute, the Food and Drug Administration, industry, and academia are needed to ensure that surrogate end point-based chemoprevention indications are feasible. Demonstration of clinical benefit on preventing IEN (as described above for FAP, sporadic colorectal adenomas, superficial bladder cancers, and Barrett’s esophagus) is one possible strategy. A second approach would follow the accelerated approval pathway for gaining marketing approval as defined in 21 Code of Federal Regulations Section 314.500. This mechanism allows early marketing approval based on scientifically strongly supported surrogate end points for disease incidence in the setting of life-threatening disease such as cancer.

References


Perspectives on Surrogate End Points in the Development of Drugs that Reduce the Risk of Cancer


Updated version  Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/9/2/127

Cited articles  This article cites 78 articles, 20 of which you can access for free at:
http://cebp.aacrjournals.org/content/9/2/127.full#ref-list-1

Citing articles  This article has been cited by 18 HighWire-hosted articles. Access the articles at:
http://cebp.aacrjournals.org/content/9/2/127.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.