Assessing Efficacy in Early-Phase Cancer Prevention Trials: The Case of Oral Premalignancy

Eva Szabo

In the current drug development paradigm, regulatory approval requires agent efficacy and safety that is demonstrated in rigorous randomized controlled clinical trials. Given the length of time and intense resource commitment that are required for phase III cancer prevention clinical trials, it is imperative that only agents that are highly likely to meet these standards be selected for late-phase clinical development. The major challenge for the field of cancer prevention (and cancer treatment) is to identify the best agents to move forward into definitive phase III testing. Preliminary indications of activity against carcinogenesis are obtained from in vitro and animal in vivo experiments (including mechanistic analyses), epidemiologic case-control and cohort studies, and early-phase clinical trials or secondary end points from clinical trials done for other indications (1). Early-phase clinical trials have the potential to be most informative because they actually test the specific agent at the requisite dose in human beings and assess a surrogate end point that should predict the phase III end point of cancer development. Identification of suitable end points for early-phase clinical trials, however, has been problematic.

Considerations for End Point Selection in Early-Phase Prevention Clinical Trials

If one of the main goals of early-phase prevention trials is to show that a drug is active enough to justify further development, then the choice of end point is one of the most critical aspects of the design of these trials. Clearly, the ideal end point would be one that is so integrally linked to cancer development that, when modulated, it would predict the definitive outcome of cancer development with 100% certainty. Unfortunately, the current understanding of epithelial cancer etiology and biology is insufficient to identify such ideal end points. Nevertheless, less-than-ideal end point markers can be very informative and help with the “go/no-go” decision following earlier-phase trials, as long as the limitations of these markers are understood.

Premalignant lesions causally linked to the development of cancer have been the “gold standard” end point for early-phase prevention trials in many organ sites (2). If the major pathway to cancer proceeds via an obligate precursor lesion, then modulation of that premalignant lesion should correlate strongly with cancer development. Colorectal adenomas represent one such precursor, and colonoscopic removal of adenomas has become standard of care because most colorectal cancers arise in previous adenomas. An accumulating body of evidence supports the hypothesis that chemoprevention of colorectal adenomas will also reduce the incidence of colorectal cancer, although this has not been formally tested in clinical trials simultaneously assessing the effect of an intervention on both adenomas and carcinomas (3–5).

Histologic premalignancy as a trial end point, however, has potential pitfalls. If not all pathways of carcinogenesis proceed through a specific histologic precursor, then an intervention that effectively modulates that precursor (but not necessarily other relevant pathways) may grossly overestimate the true effect on cancer incidence. If the intervention works at a later stage of carcinogenesis (for instance, in preventing the transition from severe dysplasia to invasive cancer) but does not suppress the histologic precursor of interest (for instance, a hyperplastic or mildly dysplastic lesion), an effective therapy could be missed. This principle is illustrated by a recently reported substudy of the Adenoma Prevention with Celecoxib trial showing that celecoxib prevented advanced colorectal adenomas but not aberrant crypt foci, thought to be an early precursor of colorectal cancer (6). Furthermore, because not all causally linked histologic precursors progress to invasive cancer, the regression of the nonprogressive-lesion subset may well lead to overestimating the true effect on cancer incidence. Therefore, it is critical to better understand the natural history of premalignant lesions and to understand how chemopreventive interventions work so that premalignancy end points can be not only appropriately paired with specific interventions but also properly interpreted. In the meantime, however, histologic end points serve as important indicators of preliminary efficacy in early-phase clinical trials.

Can molecular markers replace clinical or histologic markers as primary end points of early-phase trials? This question pertains only indirectly to the much bigger and more complex issue of validated surrogate end point biomarkers, which would be so closely linked to cancer incidence that they could replace the cancer end point of definitive prevention trials. Although early-phase molecular markers may one day prove to be useful, the answer to the question posed here does not demand such a high level of evidence. The same caveats about histologic premalignant lesions apply to molecular markers, as summarized elsewhere (7, 8). Ideally, the marker should be intrinsic to the process of carcinogenesis as well as to the mechanism of action of the intervention, such that its modulation correlates highly with the course of disease. Its expression should differ in normal and at-risk epithelium, and it should be easily and reproducibly measured from biological specimens obtained serially during clinical trials. Although sensitive to modulation by active interventions, the expression...
of the marker should have minimal spontaneous fluctuations. If molecular markers that satisfy all or most of these criteria can be identified, they could potentially be as informative as clinical or histologic premalignancy, or more so.

Measuring a pharmacodynamic end point is useful in affirming that a drug gets to its target and has its intended effect. For example, tissue polyamines (the product of ornithine decarboxylase, the target of difluoromethylornithine) were used as an early-phase trial end point to determine the optimal dose of difluoromethylornithine for a later-phase trial of this drug (combined with sulindac) in preventing colorectal adenomas (9). However, it is uninformative on true chemopreventive efficacy unless it is well established that the particular signal transduction pathway is critical to, and that inhibition of the pathway reduces, cancer progression. On the other hand, if the drug target is not modulated by the intervention, then efficacy is unlikely (unless the wrong target is being measured), and drug development should cease.

In a similar vein, dysregulated proliferation and apoptosis are well-known hallmarks of carcinogenesis and frequently are measured in chemoprevention trials, but it is not clear whether inhibition of proliferation or induction of apoptosis truly correlates with cancer development in patients (10, 11). Depending on its mechanism of action, an intervention may have desirable effects on proliferation and/or apoptosis markers and yet have no effect on cancer development or may not have desirable marker effects and may still reduce cancer development. The chances of such false-positive or false-negative indications are increased by an incomplete understanding of carcinogenesis.

If carcinogenesis in a specific organ is thoroughly understood, however, a surrogate marker can be extremely informative. For example, the link between human papillomavirus infection and subsequent cervical cancer development is well established, and the efficacy of targeting infection to prevent premalignancy and cancer has been confirmed by phase III clinical trials (12, 13). Therefore, preventing infection with high-risk serotypes was an appropriate marker and end point for early-phase clinical trials of the human papillomavirus vaccine, although this marker end point would not have been useful in trials targeting progression of disease in already infected individuals because the vaccine does not treat human papillomavirus infection. The goal, then, is to identify markers that are clinically relevant and sensitive to intervention.

**The Case of Oral Premalignancy**

In this issue of the journal, Wirth et al. (14) report a clinical trial with a molecular end point in premalignant lesions that highlights many issues pertaining to response assessments in early-phase clinical trials in general and in oral premalignancy in particular. Mindful of the caveats discussed above, how should we interpret the response assessment in the study and what can we learn? The investigators assessed the effect of celecoxib on the measurable downstream drug-effect biomarker prostaglandin E2 (PGE2) in patients with oral premalignant lesions. The accessibility to repeated measurements and biopsy, as well as a strong association with oral cancer, identifies oral premalignant lesions such as leukoplakia and erythroplakia as excellent model systems for the study of chemopreventive interventions. Oral premalignant lesions progress to squamous cell carcinoma at a rate of 17% to 36% within 8 years (15); they regress at a rate of from 10% to 30% (16, 17). In patients with oral leukoplakia, ~60% of cancers develop in the site of the original lesion, providing a rationale both for local targeting of the lesions (to extirpate the cause of 60% of cancers) and for systemic treatment (to prevent the 40% of cancers arising beyond the original lesion site; ref. 18). With this tight connection to cancer, oral leukoplakia is an appropriate end point for early-phase clinical trials, and many such trials have measured the modulation of the size and/or histology of such lesions.

Interpretation of the primary response assessment is complicated, however, even in readily accessible lesions such as oral premalignant lesions. Histologic characteristics may vary considerably within any given lesion. Although biopsies typically are done in the most clinically suspicious areas, the level of histologic abnormality in areas that are not removed is not known. Therefore, it is possible that lesions biopsied after a chemopreventive intervention were actually more (or less) histologically advanced than was indicated by the baseline biopsy. Conversely, after part or all of a lesion is removed for the baseline biopsy, it is difficult to determine whether a response in histology is due to baseline resection of the most abnormal part or to the chemopreventive intervention. This issue is highlighted in the study by Wirth et al. (14). Five participants had carcinoma in situ and underwent gross resection at baseline, leaving margins with dysplasia. It is unclear whether the decreased severity of dysplasia after celecoxib treatment was due to the drug or the resection. Only blinded, placebo-controlled trials with adequate participant numbers can sort out this issue, pointing out the importance of being careful in drawing conclusions from histologic response in small open-label trials. It also is becoming increasingly clear that accurate reporting of study methods is critical to accurate study interpretation. The issue of inter- and intra-pathologist variability in grading dysplasia will not be discussed here, but is yet another potential pitfall in assessing histologic responses.

Assessing clinical leukoplakia response (via objective measurements of visible lesions) also can be complicated. First, many leukoplakias have irregular borders and may be difficult to measure accurately. Second, biopsies can significantly reduce lesion size. The study by Wirth et al. required two adjacent 3- to 4-mm punch biopsies at baseline, leaving insufficient visible disease to monitor clinical response in several patients. Last, histologically abnormal areas may extend well beyond the clinically visible field, as highlighted by dysplasia in the resection margins of the five patients with carcinoma in situ discussed in the preceding paragraph. Poh et al. (19, 20) documented (with fluorescence visualization) occult high-risk premalignant changes after visual inspection failed to identify any abnormalities in oral mucosa. Furthermore, several groups have established that genetic abnormalities can persist in sites of oral premalignant lesions that either were resected or completely responded (clinically and histologically) to chemoprevention (21, 22).

One drawback to the leukoplakia end point is that it is not yet known whether clinical or histologic regression of leukoplakia will eventually translate into decreased cancer incidence. The only published data on this question from a prospective clinical trial showed a nonsignificant association between oral premalignant lesion response and cancer development (23). An
ongoing phase III clinical trial using erlotinib in subjects with high-risk oral leukoplakia will assess both leukoplakia response and cancer incidence within the same study (24). The relationship between oral premalignant lesion response and cancer development may differ significantly between high-risk and low-risk lesions (21, 25). For the time being, however, it is reasonable to assume that resolution (clinical, histologic, and, possibly, molecular) of oral premalignant lesions is linked with clinical benefit, somewhat analogous to the association of colorectal adenoma removal with decreased colorectal cancer incidence.

Can molecular markers perform better? The prospective clinical study of PGE2 by Wirth et al. (14) adds to nearly 30 years of prior oral premalignant lesion molecular research that has attempted to address this question with studies of, for example, retinoic acid receptor-β2, p53, and cyclin D1 (26–28). PGE2 is a primary mediator of celecoxib activity and has well-documented involvement in multiple aspects of carcinogenesis in various target organs. There are several techniques for measuring PGE2 in clinical specimens, and this diversity raises important issues of standardization and cross-study comparisons. This pilot study showed that PGE2 could be measured in clinical samples (via enzyme immunoassay on frozen tissue); the results could be used to make clinical decisions stipulated in the protocol; and PGE2 levels were modulated by celecoxib in clinical lesion specimens, all showing the feasibility of PGE2 as an early-phase end point. The use of PGE2 as the primary biomarker end point of this trial, however, raises at least three additional important issues.

First, how should response of a molecular marker be defined? In other words, how much modulation of the marker is sufficient for clinical benefit? Wirth et al. defined response as a ≥30% decrease in PGE2 (using this criterion to make treatment decisions), which relates to the 30% decrease in tumor size that defines partial response according to the widely used Response Evaluation Criteria in Solid Tumors. Although this definition seems to be reasonable, it is guided by few data, especially from clinical specimens. The definition of quantitative response criteria remains an important challenge to the use of molecular markers as primary end points in early-phase clinical trials. Statistical significance does not imply biological significance or clinical benefit.

Second, what conclusions about clinical benefit can be derived from the results of biomarker modulation trials? Wirth et al. found a lack of correlation between PGE2 modulation and dysplasia regression, although small participant numbers and potential problems with dysplasia assessment (discussed above) complicate the interpretation of this finding. Papadimitrakopoulou et al. recently found no significant differences in clinical or histologic response rates in oral premalignant lesion patients treated with celecoxib versus with placebo in a phase II trial (12). Unfortunately, this trial tested a lower dose of celecoxib (200 mg twice daily) than did the Wirth et al. study, although six participants were treated at the same high dose of 400 mg twice daily in an unblinded, nonrandomized extension of the study (only three were evaluable). Therefore, neither study is informative about the clinical benefit of high-dose celecoxib, and in the absence of a better efficacy marker to link to, the clinical benefit of PGE2 modulation is unknown. Ideally, the Wirth study would have preceded the Papadimitrakopoulou study, elucidating molecular PGE2 effects of celecoxib at 200 mg twice daily or 400 mg twice daily and thus selecting the optimal dose for the larger study with a clinical primary end point. Studies in two other settings similarly illustrate the importance of early-phase dose-finding. In the lung, celecoxib at 400 mg twice daily modulated intratumoral PGE2 expression (29, 30); in Barrett’s esophagus, celecoxib at 200 mg twice daily did not suppress premalignancy or modulate PGE2 (31).

Third, clinical benefit cannot be considered in isolation from safety, especially in the context of chemoprevention, where unexpected adverse events already have severely limited the use of agents with established preventive activity. Ideally, efficacy biomarkers one day will be paired with safety biomarkers as joint end points of early-phase chemoprevention in the oral cavity or other sites. The case of celecoxib is particularly intriguing because this has been one of the most promising chemopreventive agents, with proven efficacy, and yet is associated with serious adverse cardiovascular events (32–34), raising serious medical and social obstacles to its use in chemoprevention. Safety issues involving celecoxib in chemoprevention have been discussed in detail elsewhere (35).

Conclusions

The purpose of this perspective is to highlight issues critical to interpreting the results of early-phase clinical trials, which bear enormous weight on the decision to proceed to definitive phase III cancer prevention trials and thus have broad scientific and fiscal implications. Even early-phase investigations of oral premalignant lesions, which have the appeal of being measurable and accessible to biopsy, are challenging to interpret. Nevertheless, early-phase clinical trials with biomarker end points are interpretable, as long as the limitations of their end point assessments are understood. The conclusion from some biomarker-based studies may well be that there is insufficient information and further early-phase studies are needed before a recommendation to proceed with the next phase of drug development is warranted. Accompanied by a proper understanding of how an early-phase trial was conducted, how its end points were assessed, and the implications of the effect of intervention on its end points, appropriate decisions about proceeding to a definitive clinical trial can be based on early-phase trial results. It is the integration of clinical, histologic, and molecular information that is most likely to best inform the next phase of clinical testing.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

Assessing Efficacy in Early-Phase Cancer Prevention Trials: The Case of Oral Premalignancy

Eva Szabo


**Updated version**
Access the most recent version of this article at:
http://cancerpreventionresearch.aacrjournals.org/content/1/5/312

**Supplementary Material**
Access the most recent supplemental material at:
http://cancerpreventionresearch.aacrjournals.org/content/suppl/2008/10/14/1.5.312.DC1

**Cited articles**
This article cites 33 articles, 19 of which you can access for free at:
http://cancerpreventionresearch.aacrjournals.org/content/1/5/312.full.html#ref-list-1

**Citing articles**
This article has been cited by 3 HighWire-hosted articles. Access the articles at:
/content/1/5/312.full.html#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.