‘Validation’/’Qualification’ of Surrogate End points: A Cancer Perspective

Arthur Schatzkin, M.D., Dr.P.H.
Division of Cancer Epidemiology and Genetics
National Cancer Institute
2 Different Questions

• Q1: Does exposure-related biomarker *predict* cancer?
• Q2: Can biomarker be used as *surrogate end point* in studies of EB vs. cancer?

‘Cancer’
- Incidence: primary prevention
- Recurrence, death (prognosis): secondary prevention
Question 1 (Prediction)

Is exposure-related biomarker associated with cancer?
Relative Risk

<table>
<thead>
<tr>
<th>Cancer</th>
<th>No Cancer</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Marker +</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>Marker -</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>a + c</td>
<td>b + d</td>
<td>N</td>
</tr>
</tbody>
</table>

RR = \( \frac{a/(a+b)}{c/(c+d)} \)
### Attributable Proportion/Etiologic Fraction

<table>
<thead>
<tr>
<th></th>
<th>Cancer</th>
<th>No Cancer</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Marker +</strong></td>
<td>a</td>
<td>b</td>
<td>a + b</td>
</tr>
<tr>
<td><strong>Marker -</strong></td>
<td>c</td>
<td>d</td>
<td>c + d</td>
</tr>
<tr>
<td></td>
<td>a + c</td>
<td>b + d</td>
<td>N</td>
</tr>
</tbody>
</table>

\[
RR = \frac{a/(a+b)}{c/(c+d)} \quad \text{Sens.} = \frac{a}{a+c} \quad \text{AP} = \text{Sens} \left[\frac{1}{1-RR}\right]
\]
Question 1: (Prediction)

Why is *prediction* useful?

Helps answer this question: Is the marker on the causal pathway to cancer?

That is, can help elucidate causal pathways for exposure to cancer—i.e., ‘mechanisms’.
2 Different Questions

• Q1: Does exposure-related biomarker predict cancer?
• Q2: Can biomarker be used as surrogate end point in studies of EB vs. cancer?
Surrogate End Points

• Studies with surrogate end points can be smaller, faster, and cheaper than those with cancer outcomes

• Holds for both
  – intervention studies (trials)
  – observational epidemiologic studies
What Is ‘Validation’/’Qualification’ of Intermediate End Point Biomarkers

- Q1: the biomarker is truly on the causal pathway(s) to cancer
- Q2: the study of exposure vs. surrogate end point gives the right answer for exposure vs. cancer (a tougher requirement!)
Surrogate and Mediating Endpoints: Current Status and Future Directions
Ross L. Prentice

The identification of surrogate endpoints that can replace "true" endpoints in clinical trials could provide an important advance for the evaluation of therapeutic or preventive interventions. Outcome events that are more frequent in occurrence and more proximate in time, compared with customary disease-specific mortality or incidence outcomes, could give answers that are based on smaller trials of shorter duration. However, reliance on surrogate outcomes is justifiable only if treatment comparisons that are based on a surrogate are a faithful reflection of comparisons that are based on the true endpoint.

I took this perspective, almost 20 years ago (1), in defining a surrogate outcome to be "a response variable for which a test of the null hypothesis of no relationship to the treatment groups under comparison is also a valid test of the corresponding null hypothesis based on the true endpoint," thinking that this was a minimal requirement for a short-term outcome to provide some reliable treatment effect information for the longer-term outcome. However, this apparently simple requirement translates to some strong restrictions on the relationship of the treatment to the surrogate and true outcomes. Consider a treatment indicator variable x, a time to response surrogate S, and a time to response true outcome T. A dependence of T on x will imply a dependence of S on x if

1. the hazard rate for T depends on S and
2. the hazard rate for T given S does not depend on x.

The first criterion is typically readily verified empirically, whereas the second, which requires the surrogate to fully mediate the treatment effect on true outcome, is not. Rather, empirical data alone, even if extensive, will not provide certainty concerning criterion 2. Criterion 2 typically entails detailed knowledge of the biological pathways whereby x may affect T and detailed knowledge about the time course of such effects—knowledge that one would not expect to be available if there is uncertainty concerning whether x has any effect on T. For criterion 2 to hold, the surrogate must be comprehensive in being responsive to all pertinent pathways and the implications of the surrogate event occurrence for the true outcome risk must be equivalent in each treatment group being compared.

The article by Ray et al. (2) in this issue considers both the occurrence of distant metastases and general clinical treatment failure as potential surrogate outcomes for prostate cancer-specific death in the context of evaluating the effect of long-term androgen deprivation therapy among prostate cancer patients with locally advanced disease. Among patients who were alive 3 years after randomization, the hazard ratio for prostate cancer-specific death, when long-term androgen deprivation therapy was compared with control treatment, was 0.76 (95% confidence interval [CI]=0.51 to 1.11) among patients without distant metastases and 0.95 (95% CI=0.69 to 1.30) among patients with distant metastases. In the context of a corresponding unconditional hazard ratio of 0.69 (95% CI=0.52 to 0.93) for prostate cancer-specific mortality, these analyses convey useful information about the importance of a reduction of distant metastases in mediating the treatment effect on prostate cancer-specific survival, but they do not provide persuasive information concerning the ability of distant metastases to fulfill criterion 2. For example, the estimated prostate cancer-specific mortality rate among men without distant metastases is estimated to be 24% lower in the long-term androgen deprivation group than in the control group. Is this a chance observation or does longer term deprivation have some impact, for example, on local or regional recurrence or on the timing of distant metastasis detection?

Consideration of general clinical treatment failure as a potential surrogate (2) can be viewed as an effort to encompass pathways, in addition to reduction in distant metastases incidence, whereby the treatment may affect prostate cancer-specific death rate. General clinical treatment failure was defined as the time to first occurrence of local prostate recurrence, documented regional or distant metastasis, initiation of androgen deprivation therapy after protocol-directed treatment, or a prostate-specific antigen level of 25 ng/mL or higher after completion of radiation therapy. In spite of the stringency of criteria 1 and 2, a further criterion, as described previously (1), is needed to ensure that a dependence of S on x translates to a dependence of T on x. In this context, one can ask whether some events are included in general clinical treatment failure that help to establish an effect of x on S but have little or no implication concerning an effect of x on T. For example, the initiation of extra-protocol androgen deprivation therapy evidently may differ between treatment groups for purely artificial reasons (i.e., such therapy is unlikely in the first 2 years after randomization for men assigned to the long-term deprivation group). Also, the potential surrogate, general clinical treatment failure, may have a substantial noise component due to issues in defining local recurrence or due to patient concerns leading to unnecessary androgen deprivation therapy. Beyond these issues, the hazard ratio for long-term androgen deprivation vs control treatment was 0.88 (95% CI=0.90 to 1.63) before general clinical treatment failure and 0.81 (95% CI=0.58 to 1.14) after general clinical treatment failure, so that once again these empirical data do not provide convincing evidence.
Surrogate Validity

3 conditions needed for validity:

• Marker associated with cancer
  – Relative risk
  – Attributable Proportion

• Exposure/rx associated with marker

• Marker mediates association between exposure/rx and cancer
BMI vs. breast cancer

BMI † **estrogens** † breast ca
Surrogate Validity

3 conditions needed for validity:
• Marker associated with cancer
  – Relative risk
  – Attributable Proportion
• Exposure/rx associated with marker
• Marker mediates association between exposure/rx and cancer
## Estrogen vs. Breast Cancer (RR’s)*

<table>
<thead>
<tr>
<th></th>
<th>Estradiol</th>
<th>Free Estradiol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Q2</td>
<td>1.42</td>
<td>1.38</td>
</tr>
<tr>
<td>Q3</td>
<td>1.21</td>
<td>1.84</td>
</tr>
<tr>
<td>Q4</td>
<td>1.80</td>
<td>2.24</td>
</tr>
<tr>
<td>Q5</td>
<td>2.00</td>
<td>2.58</td>
</tr>
</tbody>
</table>

P<.001  P<.001

*JNCI 2002; 94:606-161
Surrogate Validity

3 conditions needed for validity:

• Marker associated with cancer
• Exposure associated with marker
  – RR, % change
• Marker mediates association between exposure/rx and cancer
**BMI vs. Estrogen**
(geom. mean hormone conc.)*

<table>
<thead>
<tr>
<th>BMI</th>
<th>Estradiol (pmol/L)</th>
<th>Free Estradiol (pmol/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;22.5</td>
<td>30.0</td>
<td>0.40</td>
</tr>
<tr>
<td>22.5-24.9</td>
<td>34.8</td>
<td>0.51</td>
</tr>
<tr>
<td>25.0-27.4</td>
<td>37.3</td>
<td>0.56</td>
</tr>
<tr>
<td>27.5-29.9</td>
<td>43.2</td>
<td>0.68</td>
</tr>
<tr>
<td>30.0+</td>
<td>54.9</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*JNCI 2003; 95:1218-26*
Surrogate Validity

3 conditions needed for validity:

• Marker associated with cancer
  – Relative risk
  – Attributable Proportion = Sensitivity (1-1/RR)

• Exposure/rx associated with marker
  – RR, % change

• Marker mediates association between exposure/rx and cancer
<table>
<thead>
<tr>
<th>Adjusted for free estradiol</th>
<th>RR (95% CI) for BMI (increase of 5 kg/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>1.19 (1.05-1.34)**</td>
</tr>
<tr>
<td>Yes</td>
<td>1.02 (0.89-1.17)</td>
</tr>
</tbody>
</table>

*JNCI 2003; 95:1218-26

**Prentice criterion 1
<table>
<thead>
<tr>
<th>Odds ratio</th>
<th>1</th>
<th>2</th>
<th>3–5</th>
<th>6–9</th>
<th>&gt;10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unadjusted</td>
<td>1.0</td>
<td>1.7</td>
<td>3.1*</td>
<td>4.7*</td>
<td>4.4*</td>
</tr>
<tr>
<td>Adjusted for HPV status</td>
<td>1.0</td>
<td>1.0</td>
<td>1.1</td>
<td>1.5</td>
<td>1.6</td>
</tr>
</tbody>
</table>

*p < 0.05. HPV, human papillomavirus.
‘Validation’ of Surrogate End Point Biomarkers

• Being on the causal pathway does not guarantee surrogate end point validity
Normal mucosa + E

Hyperproliferation

↓ Apoptosis
↓ Cellular adhesion factors

Neoplasia/cancer

Alternative pathway is problematic
Proliferation Markers as Surrogate End Points for Colorectal Cancer: Inferences

- Hyperproliferation may give the wrong answer about an intervention agent’s effect on CRC
  - **Agent reduces proliferation**, reduces apoptosis, has no effect on CRC
  - **Agent has no effect on proliferation**, increases apoptosis, reduces CRC
‘Validation’ of Surrogate End Point Biomarkers

- Being on the causal pathway does not guarantee surrogate end point validity
- Having a high AP/EF (meaning all or most of cancer goes through biomarker) does not guarantee surrogate end point validity
Colorectal Adenomas are Pretty Good but not Definitive Surrogates

- **Adenoma heterogeneity**: rx affects only ‘innocent’ adenomas († false positive result) or only the few ‘bad adenomas’ (‡ false null result)
Colorectal Adenomas are Pretty Good but not Definitive Surrogates

- **Timing:** in polyp trials, no information on early (pre-adenoma) and minimal information on late (small to large adenoma/cancer) events
Hematopoietic Cancer

- Benzene – ↓ WBC ‡ leukemia
- Permethrin ‡ MGUS ‡ multiple myeloma
Benzene, ↓WBC, Leukemia

- Reduced WBC associated with subsequent leukemia among benzene workers
- Strong benzene-leukemia evidence
- Regulation based on ↓WBC
Permethrin, MGUS, Multiple Myeloma

- All MM preceded by MGUS; but not all MGUS goes to MM.
- Permethrin-MM connection: limited data
- How strong can permethrin-MM inferences be based on MGUS?
More on Surrogate Validity

• A surrogate end point valid for one exposure or intervention vs. a cancer is not necessarily valid for a 2\textsuperscript{nd} exposure or intervention

• Why? Because an alternative pathway to cancer may exist
2-Stage Strategy Not a Lock

• Intervention alters surrogate marker (Stage 1) AND surrogate marker is associated with cancer (Stage 2)

• Counter-example? HRT raises HDL; HDL inversely related to CVD; but HRT does not protect vs. CVD
  – Alternative pathway(s)
Evaluating the effectiveness of the New England Journal of Medicine document, we find that the text is clearly separated into sections such as abstract, introduction, and results. The document appears to be discussing health-related research, possibly focusing on cardiovascular or medical topics. The text is structured in a standard format, making it readable and comprehensive for further analysis or study. The document layout, with clearly defined sections and organized content, aids in understanding the research being presented. The use of specific headings and subheadings, along with descriptive titles, highlights the document's focus on specific areas of medical research or clinical trials. The document is informative and appears to be aimed at professionals or researchers in the field of medicine.
3 Sobering Stories (1)

- Avandia (Rosiglitazone)
  - Lowered glycated hemoglobin level
  - Increased congestive heart failure and cardiovascular ischemia

- Folate and B-vitamins
  - Lowered homocysteine
  - Did not lower cardiovascular events
3 Sobering Stories (2)

- Torcetrapib*
  - Lowered LDL (72%), raised HDL (25%)
  - Increased cardiovascular events (25%) and all cause mortality (58%)

*torcetrapib plus atorvastatin vs. atorvastatin alone
Surrogates for Prostate Ca Survival (1)

ARTICLE

Potential Surrogate Endpoints for Prostate Cancer Survival: Analysis of a Phase III Randomized Trial

Background
The identification of surrogate endpoints for prostate cancer-specific survival may shorten the length of clinical trials for prostate cancer. We evaluated distinct metastases and general clinical treatment failure as potential surrogates for prostate cancer-specific survival by use of data from the Radiation Therapy and Oncology Group (RTOG) 92-02 randomized trial.

Methods
Patients (n = 1,051) randomly assigned and 1,031 available for this analysis) with locally advanced prostate cancer were treated with 4 months of androgen deprivation and concurrent androgen deprivation therapy with external beam radiation therapy and then randomly assigned to an additional therapy (control arm) or 24 additional months of androgen deprivation therapy (experimental arm). Data of treatment analysis at 3 and 5 years for general clinical treatment failure (defined as documented local disease progression, regional or distant metastases, initiation of androgen suppression therapy, or a prostate-specific antigen level of 25 ng/ml or higher after radiation therapy) as well as prostate-specific antigen level of 20 ng/ml or higher after radiation therapy) and/or distant metastases were tested as surrogate endpoints for prostate cancer-specific survival at 10 years by use of Kaplan-Meier survival curves. All statistical tests were two-tailed.

Results
At 3 years, 1,036 patients were alive and contributed data for analysis. Both distinct metastases and general clinical treatment failure at 3 years were consistent with all four of Pringle’s criteria for being surrogate endpoints for prostate cancer-specific survival at 10 years. At 5 years, 1,028 patients were alive and contributed data for analysis. Although prostate cancer-specific survival was not statistically significant different between treatment arms at 5 years (P = .28), both endpoints were consistent with Pringle’s survival criteria.

Conclusions
Distant metastases and general clinical treatment failure at 3 years may be considered surrogate endpoints for prostate cancer-specific survival at 10 years. These endpoints, however, must be validated in other datasets.

J Natl Cancer Inst 2006;101:285-296

Prostate cancer is a malignancy with a long survival history. In men who are initially diagnosed with high-grade and locally advanced prostate cancer often survive for many years. Because of the long survival time, clinical trials of prostate cancer that are designed with primary endpoints of overall or prostate cancer-specific survival require long follow-up periods, especially those evaluating treatments for clinically localized disease. The time required for the conception, design, conduct, analysis, and initial reporting of a prostate cancer clinical trial often approach 10 years (1-4). Identification of surrogate endpoints for prostate cancer cause-specific or overall survival would shorten the time required to conduct prostate cancer clinical trials and thus improve the chances of finding better treatments for prostate cancer. For patients with localized prostate cancer, the ideal surrogate endpoints for a survival study are clinical information that is available soon after definitive local therapy that will identify patients highly likely to die of their disease. The findings from this study would most directly apply to patients treated with primary external beam radiation therapy; further research is required to show that the findings could also apply to surgically treated patients.

The Radiation Therapy and Oncology Group (RTOG) 92-02 trial, a phase III, randomized multi-institutional clinical trial that was conducted between June 1992, and April 17, 1995, during the

Affiliations of authors: Radiation Associates of Arizona, Phoenix, AZ; MECP, Department of Radiation Oncology, University of Virginia, Charlottesville, VA; and Department of Radiation Oncology, University of Washington, Seattle, WA. J Natl Cancer Inst 2006;101:285-296.

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by the National Cancer Institute (CA-85920 and CA-66991) and by the Department of Defense-Prostate Cancer Research Program (PCRP/06-004).

See “Funding” and “Home” for information about “References.”

DOI: 10.1093/jnci/djl049

© The Authors 2006. Published by Oxford University Press. This article is a U.S. Government work and is in the public domain in the United States.
Surrogates for Prostate Ca Survival (2)

- Potential surrogates:
  - Distant metastasis
  - ‘General clinical treatment failure’

- Conclusion: these end points may be candidate surrogates for prostate cancer-specific survival at 10 years
Surrogates for Prostate Ca Survival (3)

• BUT—it is possible that a treatment effect on metastasis/rx failure does not have equivalent effect on prostate survival:
  – Rx improves surrogate end points but not survival (b/o some alternative adverse effect on mortality)
  – Rx has no impact on surrogates but improves survival (b/o some alternative beneficial effect on mortality)

~Prentice RL. JNCI 2009;101:216-7
Statistical Considerations: Error in Measurement of Biomarkers

- Measurement error will attenuate associations
  - Exposure/rx vs. marker
  - Marker vs. cancer
- Measurement error can lead to underestimation:
  - Predictive ability of biomarker
  - Extent to which surrogate mediates the effect of exposure, rx on survival
Surrogate End Point
Validity/Qualification:
Summary (1)

- **Totality** of causal connections is key
- High AP is supportive but not definitive (b/o heterogeneity, timing issues)
Surrogate End Point
Validity/Qualification:
Summary (2)

• Mediation of exposure-cancer relation supports validity of surrogacy
• Validity is more assured for surrogates both necessary for and relatively close developmentally to cancer (e.g., CIN3)
Surrogate End Point
Validity/Qualification:
Summary (3)

• Advances in validating/qualifying potential surrogate endpoints may come from:
  – Meta-analytic approaches (esp. given large sample sizes required for evaluating mediation)
  – ‘Omics’ data (transcriptomics, proteomics, metabolomics): potentially comprehensive characterization of multiple pathways
HIGHWAY OF LIFE

SLOW--IRONIC TWISTS AHEAD
Irony of Surrogate Validation

Large, long, costly studies needed for evaluation are precisely the studies surrogates were designed to replace.
Surrogate End Points in Cancer Prevention Research: The ‘No Free Lunch’ Law*

Inferential certainty is directly associated with study cost.

*or, you get what you pay for
Intermediate Biomarkers in Cancer Research: Conclusion (1)

• May be valuable in prediction—can help clarify causal pathways
Intermediate Biomarkers in Cancer Research: Conclusion (2)

• As surrogate end points
  – May be valuable in Phase II studies (those seeking a biologic effect) or observational (‘mechanism’) studies
Biomarkers in Cancer Research: Conclusion (3)

• The ‘savings’ resulting from use of surrogate end points comes at the cost of inferential certainty-- there’s no free lunch

• In conjunction with other studies (polyp trials + cohort studies of CRC), may enhance ‘probability of being right’
Biomarkers in Cancer Research: Conclusion (5)

• Replacing explicit cancer end points in Phase III clinical trials and observational studies is risky business
As informative as intermediate end point studies can be, we must not lose sight of the critical importance of observational epidemiologic studies and RCTs with incident cancer (or recurrence/mortality) end points.

(And part of that ‘importance’ is evaluating the role of potential surrogate end points.)
YESTERDAY IN THIS SPACE I PREDICTED THAT CANCER WOULD COME TO AN END. IT DID NOT, HOWEVER. I REGRET ANY INCONVENIENCE THIS MAY HAVE CAUSED.