CLINICAL BRIEFS

The Ongoing Evolution of Endpoints In Oncology

By Jack Alan McCain Jr.

Supplement to

MANAGED Care

Volume 19, No. 5
Supplement 1
May 2010

Genentech
A Member of the Roche Group

This publication is sponsored by Genentech

The Ongoing Evolution of Endpoints In Oncology

Overall survival (OS) has been the gold standard for oncology clinical trial endpoints. Over the years, however, such surrogate endpoints as objective response rate and progression-free survival have been employed because they can be reached faster and may offer important benefits in evaluating therapies.

Cancer is a leading cause of death in the United States, despite a modest reduction in the cancer death rate in recent years (Xu 2009). Five-year survival rates have improved in some cancers but not in others, spawning the development of new anticancer agents. However, the use of "hard" clinical endpoints like overall survival (OS) — the "gold standard" among endpoints in oncologic trials — or cancer-specific markers of survival can slow efforts to bring new interventions to market. For example, for a clinical trial of a prostate cancer drug employing OS or cancer-specific survival as an endpoint, the time from trial conception to initial reporting of results is about 10 years (Ray 2009).

Since the U.S. Food and Drug Administration's 1992 adoption of accelerated drug approval regulation, manufacturers often have turned to surrogate endpoints to speed the market arrival of new agents with the potential to save or extend lives. Surrogate endpoints commonly are used in the regular approval of oncologic drugs as well. Surrogate endpoints in oncology include objective response rate (ORR), progression-free survival (PFS), disease-free survival (DFS), time to progression (TTP), and patient-reported outcomes — reported individually or as parts of composite endpoints. Many surrogate endpoints record the time to the occurrence of some event. Although some surrogate endpoints may appear similar, managed care pharmacy and medical directors should understand that they have important differences when evaluating various studies using these endpoints. The purpose of this Clinical Brief is to help managed care decision makers understand the relevance of surrogate endpoints to contemporary oncology.

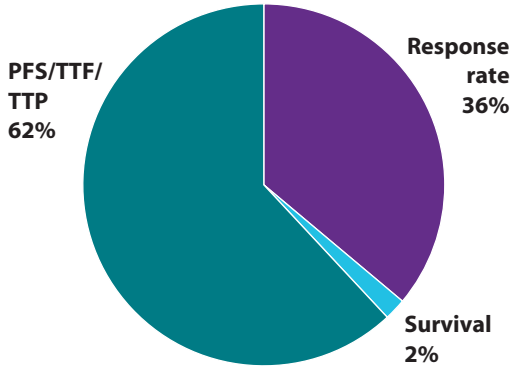
Despite its "gold standard" status, OS is increasingly regarded as a flawed measure of efficacy in colorectal cancer and advanced breast cancer (Saad 2009). Among 58 randomized clinical trials in advanced breast cancer published in eight leading medical journals during an 8-year period, OS was used as a primary endpoint only once, while surrogate time-dependent endpoints were employed 36 times (Figure 1). This reflects a trend away from OS and toward surrogate endpoints.

In comparison with overall survival, surrogate endpoints may offer numerous benefits (Table 1), the most important of which is shortening the time before patients can begin using a novel therapy. It is important for surrogates to be validated — that is, demonstrated to be adequate substitutes for overall survival. Without validation, surrogate endpoints may lead to the use of drugs that can, and in fact have, harmed patients (Shi 2009). It is insufficient for a putative surrogate endpoint only to be strongly correlated with a clinical outcome. For example,

FIGURE 1

Primary endpoints in randomized controlled trials of treatments for advanced breast cancer

In eight leading medical journals, 58 randomized clinical trials in advanced breast cancer were published from 2000 through 2007. Progression-based primary endpoints were used in a majority: progression-free survival (PFS, n=14); time to treatment failure (TTF, n=1); time to progression (TTP, n=21).



Source: Saad 2009

TABLE 1

Benefits of surrogate endpoints versus overall survival in clinical trials

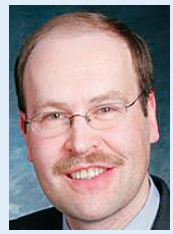
- Faster time to completion of trials
- Faster time to drug approval
- Faster receipt by patients of novel therapies
- More frequent measurements
- Insights into clinical pharmacology and mechanisms of action
- Guidance for dose selection
- More efficient screening of promising drug candidates
- Less cost to manufacturer or sponsor

Source: Shi 2009

COMMENTARY

Understanding Clinical Endpoints

A managed care analysis by Albert Tzeel, MD, MHSA, FACPE
Market Medical Officer, Great Lakes Region, Humana Inc



Albert Tzeel, MD, MHSA, FACPE

"If you don't know where you're going, you will wind up someplace else." Although most would view this quote from the master quipster Yogi Berra as humorous, ironically enough, there is an element of truth to it. And so it is when it comes to understanding surrogate endpoints used in oncology trials.

As this Clinical Brief notes, sometimes we can miss the forest for the trees. If we strictly rely on a "hard" clinical endpoint, which overall survival is, then we miss the opportunity to provide other benefits to our membership.

Surrogate endpoints provide three global benefits. First, as the Clinical Brief discusses, they help speed new treatments to market by allowing for multiple alternatives to evaluate the efficacy of new cancer therapies. This allows a greater number of treatment options to be discussed between doctor and patient when shared decision making takes place. Second, they allow for assessment of impact in situations where a cure might not be attainable but some other objective variable of importance to the member/patient might. Such an approach takes into account the potential impact of a treatment on the "quality" of a cancer patient's life. And, third, they obviate the need for large numbers of patients and protracted time intervals for analysis. This also promotes quicker acceptance and approval by the U.S. Food and Drug Administration.

So why should this be important for MCO medical and/or pharmacy directors? First of all, as we all know, the rate of turnover in a typical MCO is high. For example, one study showed that for one carrier, 50% of the members turned over after two years and only 15% of members remained after eight years (Votruba 2006). Therefore, it becomes inherent upon us to have the ability to make appropriate assessments on requested treatments because, for our members, time is of the essence.

Additionally, the Clinical Brief reviews four surrogate endpoints: Objective Response Rate, Disease-Free Survival, Progression-Free Survival, and Time to Progression. Each endpoint has a different meaning as a surrogate and each is assessed in a different manner. It is rather easy for someone to mix them up and confuse their various connotations. When such a blunder happens, the decision to provide or to exclude coverage for a particular therapy might become hampered by not paying close attention to not only what the therapy can do for the member with cancer, but, also what it is supposed to do. Not appreciating an endpoint's meaning can result in poor decisions on our part and a prolonged time to treatment for our members.

If we do not understand surrogate endpoints, we will not end up where we want to be because we will find ourselves someplace else.

References

Votruba ME, Cebul RD, Rebitzer JB, Herschman R. *Insurance turnover as an impediment to improving health care quality*. Paper presented at the Economics of Population Health: Inaugural Conference of the American Society of Health Economists, Madison, Wis. 2006.«http://www.allacademic.com/meta/p93450_index.html». Accessed Feb. 1, 2010.

Disclosure:

Albert Tzeel, MD, MHSA, FACPE, reports no financial arrangements or affiliations that may constitute a conflict of interest with his commentary.

prostate-specific antigen (PSA) is a strong prognostic factor for prostate cancer, but PSA endpoints have been found to be inadequate as surrogates for OS (Shi 2009).

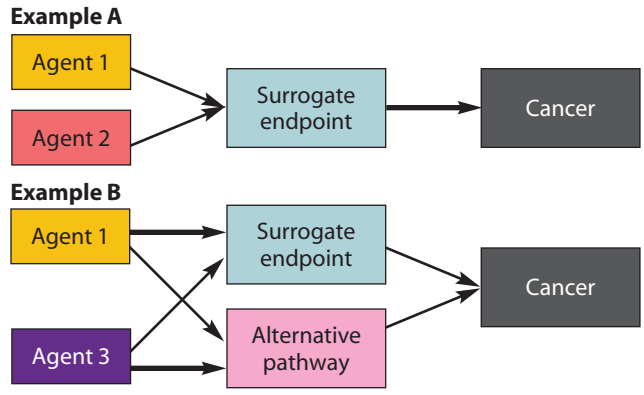
A valid surrogate endpoint must satisfy three conditions: the surrogate must be associated with the cancer; the treatment must be associated with the surrogate; and the surrogate must mediate the relationship between the treatment and the cancer (Schatzkin 2002). If these conditions are met, a valid surrogate will reliably and precisely predict the effect of the treatment on the true endpoint (Shi 2009).

Caution should be exercised in comparing results of clinical trials using surrogate endpoints or in extrapolating from those results. The fact that a surrogate has been validated for Intervention X in, say, colon cancer, does not necessarily mean the same surrogate is valid for Intervention Y in colon cancer, let alone Intervention Z in a different cancer. Each surrogate must be validated for each intervention (Figure 2, page 4), because while two interventions may affect the same surrogate, an intervention also could affect the cancer through some other mechanism — and that benefit might not be captured by the sur-

FIGURE 2

Validity of a surrogate endpoint must be established for each intervention

In example A, agents 1 and 2 exert similar effects on the surrogate endpoint and on the cancer. In example B, agents 1 and 3 exert dissimilar effects on the surrogate endpoint, and they also exert dissimilar effects on an alternative disease pathway. Hence, they exert dissimilar effects on the cancer. Relying on the surrogate endpoint to compare agents 1 and 3 could lead to the erroneous conclusion that agent 1 is superior to agent 3 because the surrogate fails to capture the effects of the agents on the alternative pathway.



Adapted from Schatzkin 2002

rogate endpoint. The same caveat applies to toxic effects that a surrogate endpoint may fail to capture. However, if a surrogate has been validated for a particular drug class, it probably is safe to assume that a new, biologically similar member of the class would affect the surrogate in a manner that is appropriately predictive of the true outcome (Schatzkin 2002). Bear in mind that not every surrogate endpoint that has been used in clinical trials of oncology drugs has been validated (Shi 2009).

Types of endpoints

Overall survival. Initially, tumor response rate (a surrogate endpoint) was sufficient by itself for FDA approval of oncologic drugs, but in the early 1980s the FDA decided response rate should not be the sole basis for approval because response rate does not adequately capture survival and clinical benefit. In 1985, the FDA stipulated that improvement in survival or patient symptoms should be demonstrated for approval of oncologic drugs (Johnson 2003). Trial designs were altered accordingly; between 1990 and 2002, 75 percent of regular and accelerated approvals for oncology therapies were based on one endpoint (Johnson 2003).

Today, the FDA still regards OS as the most reliable and preferred endpoint in a cancer trial, provided the trial can be designed to adequately assess survival (FDA 2007). Easily assessed, OS is unambiguous and is not subject to investigator bias, and it offers a clear assessment of risks and benefits.

The chief disadvantage of OS is that a long observation period is required. In addition, when patients cross over from the control arm to the experimental arm upon disease progression, the survival benefit may be obscured (FDA 2007). Second-line treatments present the risk of confounding, and as new agents come into use for second-line or refractory indications, the need grows for surrogate endpoints less likely to be confounded. If a surrogate is used, the effects of a first-line agent can be assessed before second- and third-line therapies are introduced.

Objective response rate. Because spontaneous remission is rare, ORR provides clear evidence of antitumor activity. The FDA generally defines ORR as the sum of partial responses plus complete responses (CRs) (FDA 2007). Capable of being reached within 2 or 3 months, ORR is accepted as a surrogate endpoint likely to predict demonstration of clinical benefit in randomized, controlled trials (RCTs) supporting accelerated approval (Schilsky 2002), but it also has been used in RCTs supporting regular approval (e.g., hormonal breast cancer drugs) (FDA 2007). ORR (or ORR and TTP) was the basis for 47 percent (27 of 57) of oncology drug approvals between 1990 and 2002 (Figure 3).

If a molecular target is known, an inclusion criterion can specify patients with tumors expressing the target to enrich the population with likely responders.

CR and partial responses in solid tumors often are assessed by the RECIST (Response Evaluation Criteria in Solid Tumors) methodology. RECIST provides a simplified set of criteria for evaluating tumors via an anatomical approach, using a unidimensional measure of tumor burden instead of the bidimensional measure first employed

Oncology drug approvals supported by overall survival as an endpoint

- Bevacizumab (metastatic colorectal cancer)
- Bevacizumab (unresectable nonsquamous nonsmall cell lung cancer)
- Capecitabine (metastatic breast cancer, with docetaxel after failed anthracycline treatment)
- Docetaxel (metastatic breast cancer, 2nd-line)
- Irinotecan (colon cancer, 1st-line)
- Irinotecan (colon cancer, 2nd-line)
- Leucovorin (metastatic colon cancer, in combination with fluoruracil)
- Paclitaxel (ovarian cancer, 1st-line)

by WHO. RECIST was published in 2000 (Therasse 2000) and revised in 2009 as version 1.1 (Eisenhauer 2009). Version 1.1 requires measuring a maximum of 5 lesions (down from 10 in the initial version) and a maximum of two lesions per organ (down from five). A complete response is defined by RECIST 1.1 as the disappearance of all target lesions, along with a reduction to <10 mm in the short axis of any pathological lymph node. A partial response requires a decrease of at least 30 percent in the sum of diameters of target lesions from their sum at baseline (Eisenhauer 2009).

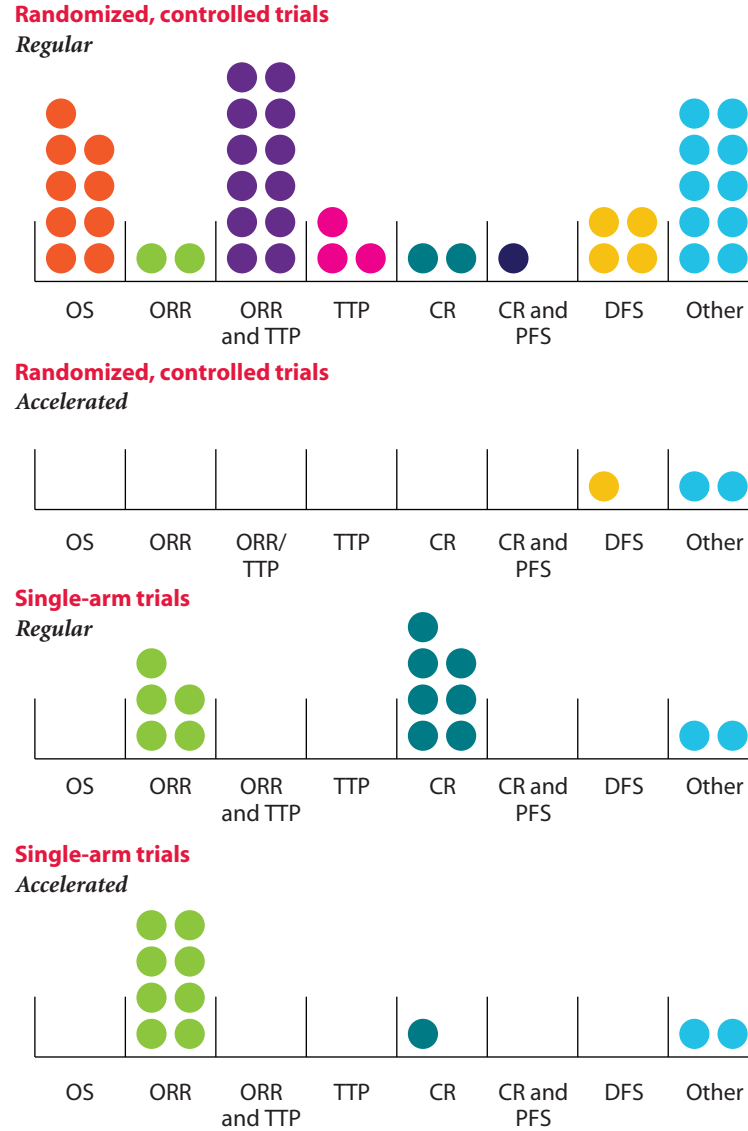
A disadvantage of ORR is that it is not a direct measure of clinical benefit, and it probably fails to capture the full benefit of treatment (Johnson 2003). Response *duration* is as important as response rate to patients, but response duration can only be estimated.

Anatomy-based tumor response criteria generally reflect the limitations of imaging technology that were in place some 30 years ago when anatomy-based criteria were developed (Zhao 2009). In particular, unidimensional and bidimensional anatomical measures suffice for assessing tumor volume only when tumors are spherical initially and in response to treatment. Anatomical measurements fail to capture changes in tumor density, and they cannot distinguish viable tumor from dead tumor. Further, intra- and inter-observer variability in interpretation of RECIST data adds to uncertainty about measurements, requiring more patients and more time, or both, to establish differences (Petrick 2008). Finally, because cut-off values are crude and arbitrary, RECIST values for partial response and disease progression may be insufficient for timely detection of disease progression and regression. These shortcomings may be addressed by emerging assessments of tumor responses that capitalize on modern technology, such as measurement of tumor density, measurement of tumor volume, and quantitative functional molecular imaging that tracks biochemical processes.

FIGURE 3

Endpoints used for basis of oncology drug approvals, 1990–2002

Each circle represents an endpoint used to support marketing applications for oncology drugs approved by the FDA from Jan. 1, 1990 to Nov. 1, 2002. Many of the 57 applications reported more than one endpoint; hence, the number of circles is greater than 57.



CR=complete response, DFS=disease-free survival, ORR=objective response rate, OS=overall survival, PFS=progression-free survival, TTP=time to progression. Source: Johnson 2003

Oncology drug approvals supported by objective response rate as an endpoint

- Altretamine (refractory ovarian cancer)
- Anastrozole (breast cancer, 2nd-line; 1st-line)
- Capecitabine (refractory breast cancer)
- Exemestane (breast cancer, 2nd-line)
- Imatinib (gastrointestinal stromal tumor)
- Irinotecan (colon cancer, 2nd-line)
- Letrozole (breast cancer, 2nd-line; 1st-line)

Disease-free survival. DFS (also known as recurrence-free survival) is a composite endpoint combining time to disease recurrence and OS. The inclusion of OS addresses

Oncology drug approvals supported by disease-free survival as an endpoint

- Anastrozole (breast cancer, adjuvant therapy for postmenopausal patients with estrogen receptor-positive tumors)
- Letrozole (early breast cancer, adjuvant treatment in postmenopausal women following adjuvant tamoxifen)
- Exemestane (early breast cancer, adjuvant treatment following adjuvant tamoxifen)
- Oxaliplatin (colon cancer, adjuvant treatment of stage III)
- Tamoxifen (breast cancer, adjuvant therapy for node-negative)
- Trastuzumab (HER2-overexpressing breast cancer)

long-term adverse effects, such as toxicity or secondary malignancies (Gill 2006). DFS is useful when prolonged survival makes measurement of OS impractical (FDA 2007). For example, DFS was the primary endpoint in the Multicenter International Study of Oxaliplatin/5-Fluorouracil/Leucovorin in the Adjuvant Treatment of Colon Cancer (MOSAIC). After a median follow-up of 38 months, there was no statistically significant difference in OS between the two study groups, but in the group in which oxaliplatin was added, the hazard ratio (HR) for recurrence was 0.77 (P=.002) (André 2004). In other words, using OS as an endpoint would have suggested that oxaliplatin provided no benefit after 3 years of follow-up, but using DFS as a surrogate suggested otherwise. At 6 years, a statistically significant benefit from adding oxaliplatin to the regimen could be seen in OS, a secondary endpoint, as its HR was 0.84 (P=.046) (André 2009).

In 1990, DFS was accepted by the FDA to support the approval of tamoxifen as adjuvant therapy for node-negative breast cancer. DFS also has supported approval of adjuvant cytotoxic breast cancer therapy and adjuvant colon cancer therapy (FDA 2007). For adjuvant breast cancer treatments, DFS has become the primary basis for FDA approval (Cortazar 2009).

DFS is a strong surrogate for OS in adjuvant colon cancer (Shi 2009), but it has not been validated as a surrogate for survival in all settings, such as lung cancer (FDA 2007). Like PFS, it is subject to assessment bias, especially in open-label trials, and its definitions vary. The FDA has called for a standardization of DFS definition in adjuvant breast cancer trials to facilitate interpretation, analysis, and comparisons of trials (Cortazar 2009).

Progression-free survival. PFS is another surrogate for OS. In 1991, PFS was accepted by the FDA as an endpoint supporting regular approval of carboplatin for treatment of ovarian cancer.

PFS is a composite endpoint that is similar to DFS, combining time to disease progression with OS. It includes all deaths but allows the enrollment of fewer subjects and requires shorter follow-up than OS studies, includes measurement of stable disease, and tends to yield objective and quantitative assessments (FDA 2007).

The FDA regards PFS as a more accurate assessment of stable disease than ORR, and it favors PFS over TTP as a regulatory endpoint. Unlike ORR, PFS captures the effects of therapies that do not affect tumor size; unlike TTP, PFS honors the intent-to-treat principle by capturing death, thereby avoiding the problem of bias associated with the censoring of death in TTP analyses (see Kaplan-Meier box, page 7).

Recent interest in the adoption of progression-based endpoints (Figure 1, page 2) may be a reflection of current research interest, as endpoints often reflect the type of cancer being studied. PFS, for instance, could be expected to be used in trials of therapies for metastatic cancers, where the goal of treatment is to prolong survival. In hundreds of trials in patients with advanced breast cancer, gains in OS have been rare; hence, disease progression increasingly is used as a primary endpoint (Smith 2006). Considering the FDA's preference for DFS over TTP, DFS may continue to draw attention from investigators studying treatments for patients with metastatic cancers.

The drawbacks of PFS are that it requires frequent radiologic or other assessment at periodic intervals, which introduces imprecision into the measure because the exact time of progression usually is unknown; its definitions vary; it is subject to assessment bias, especially in open-label studies; and it has not been validated in all settings. However, it has been argued that in some of these settings the sequential use of cancer drugs with improved PFS and a favorable risk-benefit profile has led to cumulative improvement in OS such that clinically meaningful PFS improvement in high-quality trials should be regarded as evidence of clinical benefit regardless of an individual drug's effect on OS (Zhuang 2009).

Time to progression. TTP differs from PFS by censoring deaths that occur before progression. Unlike OS, TTP presents no danger of confounding from crossover or second-line therapy, and it reaches the endpoint sooner (Schilsky 2002). The endpoint must be precisely defined. In the previously mentioned study of RCTs in advanced

Oncology drug approvals supported by progression-free survival as an endpoint

- Alemtuzumab (B-cell chronic lymphocytic leukemia)
- Bevacizumab (metastatic breast cancer)
- Bevacizumab (metastatic renal cell carcinoma)
- Gemcitabine (advanced ovarian cancer, relapsed)
- Ixabepilone (metastatic or locally advanced breast cancer)
- Panitumumab (advanced colorectal cancer)
- Sorafenib (advanced renal cell carcinoma)
- Sunitinib (advanced renal cell carcinoma)

Oncology drug approvals supported by time to progression as an endpoint

- Anastrozole (breast cancer, 2nd-line)
- Anastrozole (breast cancer, 1st-line)
- Lapatinib (metastatic breast cancer overexpressing HER2)
- Topotecan (ovarian cancer, 2nd-line)
- Fulvestrant (breast cancer, 2nd-line)
- Sunitinib (gastrointestinal stromal tumor)
- Trastuzumab (metastatic breast cancer, 1st line)

breast cancer published in leading journals, however, the definition of TTP was not reported in 5 of the 21 trials using it as the primary endpoint, and in two trials TTP was used interchangeably with PFS (Saad 2009). Accepted TTP endpoints include cancer-related death, increased size of existing lesions, and emergence of new lesions (Schilsky 2002).

At best, TTP provides an estimate, because progression occurs at some unknown point between scheduled observations. It follows that the less frequent the observations, the longer the TTP. TTP is expensive because of the need to assess all possible disease sites at baseline and at each follow-up, using the same technology each time, and because it requires an RCT design to control for tumor progression occurring in the absence of treatment effect (Schilsky 2002). This endpoint also is hindered by uncertainty over the correlation between improvement in TTP and benefit for patients.

The FDA regards TTP as a more accurate assessment of stable disease than ORR (FDA 2007). A meta-analysis has suggested that in advanced breast cancer, TTP is a better surrogate for OS than tumor response and progressive disease, in women receiving first-line combination anthracycline chemotherapy (Hackshaw 2005).

In 1994, TTP supported regular approval paclitaxel as

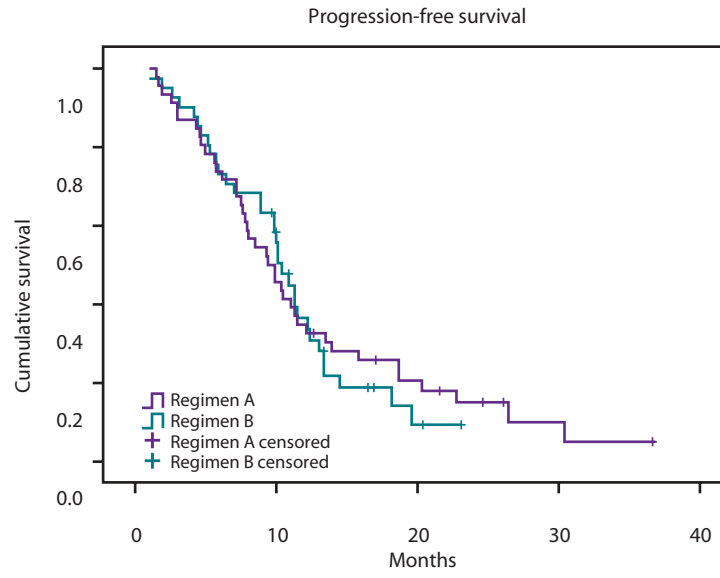
Kaplan-Meier survival curves

A Kaplan-Meier (KM) survival "curve" is a graphical means of plotting the probability of surviving until a given time without experiencing some event of interest (e.g., death, disease progression), beginning from a defined point, such as diagnosis or initiation of therapy. KM curves have a stepped appearance, with the vertical drop at the end of a time interval depicting the reduction in survival to the next interval. For various reasons, the observation period often ends before all subjects have experienced the event of interest; such patients are said to be censored. When patients are censored, it is not known when or whether the event of interest will occur, but their data are taken into account in the analysis. If patients are censored, they are shown as a small integer or a graphic device (e.g., vertical tick) above the horizontal line depicting the interval when the censoring occurred.

KM curves are built around three assumptions: (1) survival prospects for censored patients are the same as those for patients who continue to be followed; (2) survival probabilities are the same for patients recruited early and late in the study; (3) the event happens at the time specified (Bland 1998). Assumption (2) is particularly important in lengthy oncology trials because of the likelihood that treatment regimens will change over time, owing to the introduction of new therapies.

Kaplan-Meier curves depicting progression-free survival

Example shows survival rates comparing two treatment regimens



Source: Skof 2009

The log-rank statistical hypothesis test often is used to determine whether there is any difference in survival times between two KM curves, but it cannot take multiple variables into account in order to separate treatment effects from the effects of other variables affecting survival. That is accomplished with the Cox proportion hazards model. In a Cox model, the measure of risk calculated for each variable is known as the hazard ratio (HR). An HR >1 indicates increased risk for subjects with that characteristic while an HR <1 indicates decreased risk.

Timeline — Endpoints in Oncology

1950s — Physician's subjective evaluation of tumor response is endpoint initially used in clinical trials in solid tumors; until the early 1980s, tumor response rate alone suffices for FDA approval of oncology drugs (Johnson 2003)

1977 — Breast cancer specialists publish principles for evaluating response to treatment (Hayward 1977)

1980s — FDA determines that approval of cancer drugs should be based on more evidence of clinical benefit (improvement in survival, quality of life, physical functioning, or tumor-related symptoms), which is not necessarily predicted by or correlated with objective response rate (FDA 2007)

1981 — World Health Organization (WHO) adopts breast cancer criteria for evaluation of all solid tumors (Miller 1981)

1983 — Term *surrogate endpoint* first used, in context of rheumatoid arthritis research

1985 — FDA requires improvement in survival or patient symptoms for approval of oncologic drugs, deeming response rate (RR) insufficient for demonstrating survival or clinical benefit (Johnson 2003)

1989 —*Statistics in Medicine* publishes four papers establishing formal foundation for surrogate endpoints

1990s — Lack of detail in original WHO criteria leads groups to develop their own versions; use of different versions leads to difficulty in comparing results

1990 — RR is accepted by FDA as endpoint supporting regular approval of altretamine for refractory ovarian cancer (Johnson 2003)

1990 — Disease-free survival (DFS) is accepted by FDA as endpoint supporting regular approval of tamoxifen for estrogen receptor-positive breast cancers

1991 — Progression-free survival (PFS) is accepted by FDA as endpoint supporting regular approval of carboplatin for ovarian cancer

1991 — FDA and National Cancer Institute (NCI) propose DFS as endpoint in adjuvant setting where most recurrences are symptomatic (Johnson 2003)

1992 — *Accelerated approval* is added to New Drug Application regulations for investigational agents intended to treat serious or life-threatening diseases; FDA approval of a drug may be granted on basis of surrogate endpoint that is reasonably likely to predict clinical benefit

1995 — Oncology interventions begin to be approved under the accelerated approval process

1996 — Representatives of NCI, National Cancer Institute of Canada, and European Organization for Research and Treatment of Cancer begin update of WHO criteria to accommodate advances in imaging, emphasis on response rate, and rapid increase of number of investigational agents

1999 — Standardized response criteria are published for specific cancers: non-Hodgkin's lymphomas (Cheson 1999) and androgen-independent prostate cancer (Bubley 1999)

2000 — Comprehensive revision of WHO criteria is published as RECIST (Response Evaluation Criteria In Solid Tumors) (Therasse 2000) and is rapidly adopted by regulators, pharmaceutical companies, and researchers; used primarily in phase 2 trials to evaluate RR of a new agent or new combination of agents

2003 — Oncologic Drugs Advisory Committee (ODAC) recommends PFS as endpoint for advanced non-small cell lung carcinoma trials

2004 — ODAC recommends DFS as an acceptable endpoint for colon cancer drugs in surgical adjuvant setting

2006 — Guidelines published regarding use of ^{18}F-FDG PET as biomarker to assess treatment response (Shankar 2006)

2007 — FDA issues guidance regarding clinical trial endpoints used in the approval of drugs intended to treat patients with existing cancer (FDA 2007)

2009 — Version 1.1 of RECIST is published (Eisenhauer 2009)

second-line treatment of breast cancer. Since then it has supported approval of several aromatase inhibitors as treatment of breast cancer and the approval of gemcitabine for treatment of non-small cell lung cancer, among other approvals.

Patient-reported outcomes. Patient-reported outcomes, such as quality of life (QOL), complement information from traditional endpoints, generating the patient's global assessment of the direct clinical benefit of a drug (FDA 2007, Schilsky 2002). These outcomes are difficult to use, however, because they rely on subjective reporting and require double-blind placebo-controlled studies to minimize bias.

Conclusion

Although OS is the gold standard among endpoints in oncology, it can require large numbers of patients and lengthy follow-up to demonstrate efficacy. Surrogate endpoints can speed the time necessary to bring new anticancer agents to market, though bear in mind when evaluating reported results from trials that employ surrogate endpoints that each surrogate has its advantages and disadvantages.

Considerable variation has been found in the definitions of surrogate endpoints such as TTP and DFS, which hinders the comparison of trials. It also is important to remember that the validity of a surrogate endpoint should be established for each intervention. In some settings, however, high-quality trials that use PFS as an endpoint and demonstrate a favorable risk-benefit profile for an individual drug may be regarded as evidence of clinical benefit regardless of the drug's effect on OS.

Endpoints Currently in Use

Endpoint	Regulatory evidence	Study design	Advantages	Disadvantages	Drug (cancer type) approval supported with this endpoint
Overall survival	Clinical benefit for regular approval	• Randomized studies essential • Blinding not essential	• Universally accepted direct measure of benefit • Easily measured • Precisely measured	• May involve larger studies • May be affected by crossover therapy and sequential therapy • Includes noncancer deaths	Bevacizumab (mCRC) Capecitabine (colon) Irinotecan (colon) Leucovorin (colon)
Symptom endpoints (patient-reported outcomes)	Clinical benefit for regular approval	• Randomized blinded studies	• Patient perspective of direct clinical benefit	• Blinding often is difficult • Data frequently are missing or incomplete • Clinical significance of small changes is unknown • Multiple analyses • Lack of validated instruments	
Disease-free survival	Surrogate for accelerated approval or regular approval*	• Randomized studies essential • Blinding preferred • Blinded review recommended	• Smaller sample size and shorter follow-up necessary compared with survival studies	• Not statistically validated as surrogate for survival in all settings • Not precisely measured subject to assessment bias, particularly in open-label studies • Definitions vary among studies	Anastrozole (breast) Letrozole (breast) Tamoxifen (breast) Trastuzumab (breast)
Objective response rate	Surrogate for accelerated approval or regular approval*	• Single-arm or randomized studies can be used • Blinding preferred in comparative studies • Blinded review recommended	• Can be assessed in single-arm studies • Assessed earlier and in smaller studies compared with survival studies • Effect attributable to drug, not natural history	• Not a direct measure of benefit • Not a comprehensive measure of drug activity • Only a subset of patients who benefit	Anastrozole (breast) Capecitabine (breast) Exemestane (breast) Imatinib (GIST) Irinotecan (colon) Letrozole (breast)
Complete response	Surrogate for accelerated approval or regular approval*	• Single-arm or randomized studies can be used • Blinding preferred in comparative studies • Blinded review recommended	• Can be assessed in single-arm studies • Durable complete responses can represent clinical benefit • Assessed earlier and in smaller studies compared with survival studies	• Not a direct measure of benefit in all cases • Not a comprehensive measure of drug activity • Small subset of patients with benefit	Arsenic trioxide (APL) Gemtuzumab (AML) Idarubicin (AML) Trefinoin (APL)
Progression-free survival (includes all deaths) or **Time to progression** (deaths before progression censored)	Surrogate for accelerated approval or regular approval*	• Randomized studies essential • Blinding preferred • Blinded review recommended	• Smaller sample size and shorter follow-up necessary compared with survival studies • Measurement of stable disease included • Not affected by crossover or subsequent therapies • Generally based on objective and quantitative assessment	• Not statistically validated as surrogate for survival in all settings • Not precisely measured subject to assessment bias particularly in open-label studies • Definitions vary among studies • Frequent radiological or other assessments • Involves balanced timing of assessments among treatment arms	Alemtuzumab (B-CLL) Bevacizumab (mCRC) Bevacizumab (mRCC) Bevacizumab (MBC) Sorafenib (mRCC) Sunitinib (mRCC) Oxaliplatin (colon) Topotecan (ovarian) Fulvestrant (breast) Sunitinib (GIST)

*Adequacy as a surrogate endpoint for accelerated approval or regular approval is highly dependent upon other factors such as effect size, effect duration, and benefits of other available therapy.

AML=acute myelogenous leukemia, APL=acute promyelocytic leukemia, B-CLL=B-cell chronic lymphocytic leukemia , GIST=gastrointestinal stromal tumor, MBC=metastatic breast cancer, mCRC=metastatic colorectal cancer, mRCC=metastatic renal cell carcinoma.

Sources: FDA 2007, manufacturers' prescribing information

Glossary

Accelerated approval — Regulatory pathway for rapidly bringing to market a new drug intended to treat patients with serious or life-threatening illnesses; a *surrogate endpoint* can be used when accelerated approval is sought. First used in development of drugs to treat AIDS. Once accelerated approval is granted, postmarketing studies are required to demonstrate clinical benefit.

Active control — Widely used instead of placebo control in oncology trials when a therapy known to reduce mortality or serious morbidity already exists.

Biomarker — An indicator of normal biological processes, pathogenic progresses, or pharmacologic responses to a therapeutic intervention that can be objectively measured. Surrogate endpoints such as objective response rate, progression-free survival, and time-to-tumor progression are regarded as biomarkers (Fleming 2009). In some instances, surrogate endpoints have been the sole basis for FDA approval of oncology drugs. Other biomarkers provide early information of a therapeutic response, which may or may not result in improved outcomes. (Rudin 2009) These are generally not a primary endpoint in oncology regulatory trials but has been accepted by FDA as part of a composite endpoint.

Complete response (CR) — Per RECIST 1.1, the disappearance of all target lesions, along with a reduction to <10 mm in short axis of any pathological lymph node (Eisenhauer 2009). Also see *partial response.*

Disease-free survival (DFS) — A surrogate endpoint for overall survival; also known as *recurrence-free survival.* Definitions vary, but often is defined as time from randomization to recurrence or relapse, second cancer, or death. Requires randomized study, preferably blinded; blinded review is recommended (FDA 2007).

Hazard ratio — A summary of the difference between two survival curves, representing the reduction in the risk of death on treatment compared to the control. It is also used to calculate the degree of improvement in the trial's endpoint (e.g., overall survival,progression-free survival over the course of the entire study.

Noninferiority (NI) trial — An active-controlled trial showing that a new drug is not less effective than a standard regimen by a prespecified amount known as the *noninferiority margin.* In a given NI trial, it is assumed that the active control had an effect (even though it may not have), and it is further assumed that if a placebo arm had been included, placebo would have been inferior to the active control (Temple 2000).

Objective response rate (ORR) — A surrogate endpoint for *overall survival.* Defined as proportion of patients with reduction in tumor size by a predefined amount (using standardized criteria, such as RECIST) and for a minimum duration, usually measured from time of initial response to documented tumor progression; sum of partial and complete responses. Directly attributable to drug effect; often used in *single-arm trials* in refractory cancer (FDA 2007).

Overall survival (OS) — Defined as time from randomization until death from any cause; most commonly used endpoint in phase 3 trials and trials for regulatory approval. Requires randomized trial with lengthy follow-up. Can be affected by subsequent therapies. Median OS is a common endpoint in metastatic colorectal cancer (Tang 2007).

Partial response — Per RECIST 1.1, a decrease of at least 30 percent in the sum of diameters of target lesions from their sum at baseline (Eisenhauer 2009); per the World Health Organization (WHO), a decrease of 50 percent in the entire tumor burden. Also see *complete response.*

Progression-free survival (PFS) — A surrogate endpoint for *overall survival.* Defined as time from randomization to objective tumor progression or death. Unlike *time to progression* (TTP), PFS includes death from any cause as well as progression. Like TTP, it is unaffected by subsequent therapies. FDA prefers PFS rather than TTP as regulatory endpoint (FDA 2007). Note that no standard regulatory criteria define progression; *RECIST* criteria are among those used by applicants.

Progressive disease — Per RECIST 1.1, an increase of 20 percent or more in sum of diameters of target lesions relative to the smallest sum at any point during the study, along with an absolute increase of 5 mm or more in the sum; appearance of 1 or more new lesions also constitutes progression (Eisenhauer 2009); per WHO, an increase of at least 25 percent in a specific tumor lesion or in the global tumor load in one organ. Also see *stable disease.*

RECIST (Response Evaluation Criteria in Solid Tumors) — Simplified set of criteria for evaluating solid tumors, via an anatomical approach. Uses unidimensional measure of tumor burden instead of bidimensional measure first employed by WHO. First published in 2000 (Therasse 2000) and revised in 2009 as version 1.1 (Eisenhauer 2009). Version 1.1 requires measuring a maximum of 5 lesions in total (down from 10 in initial version) and maximum of 2 lesions per organ (down from 5).

Single-arm trial — A trial lacking a control arm. Useful for trials in refractory tumors for which no available therapy exists. ORR is an endpoint in single-arm trials.

Stable disease — Insufficient change to qualify as either *partial response* or *progressive disease* per RECIST 1.1 (Eisenhauer 2009) or WHO criteria.

Surrogate endpoint — An alternative endpoint, such as a biological marker, physical sign, or precursor event, that, if validated, allows conclusions to be made about the effect of an intervention on a true endpoint (i.e., a clinically meaningful endpoint) without actually observing the true endpoint (Baker 2006). When used for accelerated approval, surrogate endpoints must be reasonably likely to predict clinical benefit as well as provide a benefit over available therapy; ORR is surrogate endpoint most commonly used in accelerated approval (FDA 2007).

Time-to-event trials — Clinical trials whose endpoints measure the time until a certain event occurs, such as death, discontinuation, or tumor progression. Randomized trials are required when these endpoints are used (FDA 2007).

Time to [tumor] progression (TTP) — A surrogate endpoint for *overall survival.* Defined as time from randomization until objective tumor progression. Unlike PFS, it does not include deaths, but if most deaths are not cancer-related TTP can be acceptable endpoint. Like PFS, it is unaffected by subsequent therapies.

Time to progression of cancer symptoms — Resembles TTP but is a direct measure of clinical benefit, not a surrogate endpoint.

Time to treatment failure (TTF) — A composite endpoint measuring time from randomization to treatment discontinuation for any reason (disease progression, treatment toxicity, death). Not recommended as regulatory endpoint because it fails to clearly distinguish efficacy from toxicity, intolerance, and withdrawal (FDA 2007).

References

André T, Boni C, Mounedji-Boudiaf L, et al; Multicenter International Study of Oxaliplatin/5-Fluorouracil/Leucovorin in the Adjuvant Treatment of Colon Cancer (MOSAIC) Investigators. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. *N Engl J Med.* 2004;350:2343–2451.

André T, Boni C, Navarro M, et al. Improved overall survival with oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial. *J Clin Oncol.* 2009;27:3109–3116.

Baker SG. Surrogate endpoints: wishful thinking or reality? *J Natl Cancer Inst.* 2006;98:502–503.

Bland JM, Altman DG. Survival probabilities (the Kaplan-Meier method). *BMJ.* 1998;317:1572.

Bubley GJ, Carducci M, Dahut W, et al. Eligibility and response guidelines for phase II clinical trials in androgen-independent prostate cancer: recommendations from the Prostate-Specific Antigen Working Group. *J Clin Oncol.* 1999;17:3461–3467. Erratum: *J Clin Oncol.* 2000;18:2644. *J Clin Oncol.* 2007;25:154.

Cheson BD, Horning SJ, Coiffier B, et al. Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas. NCI Sponsored International Working Group. *J Clin Oncol.* 1999;17:1244. Erratum: *J Clin Oncol* 2000;18:351.

Cortazar P, Johnson J, Justice R, Pazdur R. Adjuvant breast cancer : FDA approval overview. *J Clin Oncol.* 2009;27(suppl). Abstract e11529.

Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer.* 2009;45:228–247.

FDA (U.S. Food and Drug Administration). *Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics.* May 2007. «http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071590.pdf». Accessed Aug. 26, 2009.

Fleming TR, Rothman MD, Lu HL. Issues in using progression-free survival when evaluating oncology products. *J Clin Oncol.* 2009;27:2874–2880.

Gill S, Sargent D. End points for adjuvant therapy trials: has the time come to accept disease-free survival as a surrogate end point for overall survival? *Oncologist.* 2006:11:624–629.

Hackshaw A, Knight A, Barrett-Lee P, Leonard R. Surrogate markers and survival in women receiving first-line combination anthracycline chemotherapy for advanced breast cancer. *Br J Cancer.* 2005;93:1215–1221.

Hayward JL, Carbone PP, Heuson JC, et al. Assessment of response to therapy in advanced breast cancer: a project of the Programme on Clinical Oncology of the International Union Against Cancer, Geneva, Switzerland. *Cancer.* 1977;39:1289–1294.

Johnson JR, Williams G, Pazdur R. End points and United States Food and Drug Administration approval of oncology drugs. *J Clin Oncol.* 2003;21:1404–1411.

Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer.* 1981;47:207–214.

Petrick N, Brown DG, Suleiman O, Myers KJ. Imaging as a tumor biomarker in oncology drug trials for lung cancer: the FDA perspective. *Clin Pharmacol Ther.* 2008;84:523–525.

Ray ME, Bae K, Hussain MH, et al. Potential surrogate endpoints for prostate cancer survival: analysis of a phase III randomized trial. *J Natl Cancer Inst.* 2009;101:228–236.

Rudin M. Noninvasive structural, functional, and molecular imaging in drug development. *Curr Opin Chem Biol.* 2009;13:360–371.

Saad ED, Katz A. Progression-free survival and time to progression as primary end points in advanced breast cancer: often used, sometimes loosely defined. *Ann Oncol.* 2009;20:460–464.

Schatzkin A, Gail M. The promise and peril of surrogate end points in cancer research. *Nat Rev Cancer.* 2002;2:19–27.

Schilsky RL. End points in cancer clinical trials and the drug approval process. *Clin Cancer Res.* 2002;8:935–938.

Shankar LK, Hoffman JM, Bacharach S, et al; National Cancer Institute. Consensus recommendations for the use of ^{18}F-FDG PET as an indicator of therapeutic response in patients in National Cancer Institute Trials. *J Nucl Med.* 2006;47:1059–1066.

Shi Q, Sargent DJ. Meta-analysis for the evaluation of surrogate endpoints in cancer clinical trials. *Int J Clin Oncol.* 2009;14:102–111.

Smith I. Goals of treatment for patients with metastatic breast cancer. *Semin Oncol.* 2006;33(1 suppl 2):S2–S5.

Tang PA, Bentzen SM, Chen EX, Siu LL. Surrogate end points for median overall survival in metastatic colorectal cancer: literature-based analysis from 39 randomized controlled trials of first-line chemotherapy. *J Clin Oncol.* 2007;25:4562–4568.

Temple R, Ellenberg SS. Placebo-controlled trials and active-control trials in the evaluation of new treatments. Part 1: ethical and scientific issues. *Ann Intern Med.* 2000;133:455–463.

Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst.* 2000;92:205–216.

Xu J, Kochanek KD, Tejada-Vera B. Deaths: preliminary data for 2007. *National Vital Statistics Reports.* http://www.cdc.gov/nchs/data/nvsr/nvsr58/nvsr58_01.pdf. Accessed Aug. 25, 2009.

Zhao B, Schwartz LH, Larson SM. Imaging surrogates of tumor response to therapy: anatomic and functional biomarkers. *J Nucl Med.* 2009;50:239–249.

Zhuang SH, Xiu L, Elsayed YA. Overall survival: a gold standard in search of a surrogate: the value of progression-free survival and time to progression as end points of drug efficacy. *Cancer J.* 2009;15:395–400.

About the Author

Jack Alan McCain Jr. is a freelance medical writer and editor based in Durham, Conn.

Disclosure: Jack Alan McCain Jr. reports no financial arrangements or affiliations that may constitute a conflict of interest.