

Oncology Endpoints in a Changing Landscape

Supplement to

MANAGED

Care

January 2016

Brought to you by

Genentech

A Member of the Roche Group

Oncology Endpoints in a Changing Landscape

INTRODUCTION

This article is an overview of the endpoints commonly used in oncology clinical trials. It is intended to serve as a high-level survey of the topic that will familiarize the reader with endpoints overall and provide context on why oncology trials today are designed with a growing variety of endpoints. This article discusses the strengths and weaknesses of commonly used endpoints, as well as emerging endpoints that reflect changes in the course of the disease and new treatment options. (See Table 1, page 9, for a summary.) It begins with some general observations about endpoints and then discusses individual endpoints in more detail.

Categories of Endpoints

Endpoints used in oncology trials can be grouped into 2 general categories: **patient-centered endpoints** and **tumor-centered endpoints** (Fiteni 2014). Patient-centered endpoints include overall survival (OS) and health-related quality of life (HR-QOL). Tumor-centered endpoints are used as surrogates for the patient-centered endpoints in clinical trials

(Fiteni 2014). Each oncology endpoint has advantages and disadvantages that vary with the type of tumor, line of treatment, disease progression, and expected survival.

OS has long been considered the “gold standard.” But when patients with cancer survive longer—as is increasingly the case because of advances in scientific knowledge and improved patient outcomes—OS becomes more difficult to measure (Wilson 2015b). OS results may be affected by patients crossing over from a control treatment to a new therapy during a clinical trial. Keeping patients on a study treatment when other potentially efficacious drugs are available has ethical implications. Generally, as patients survive longer, it becomes difficult to isolate the treatment effects of a single drug because of the multiple therapeutic options that are available.

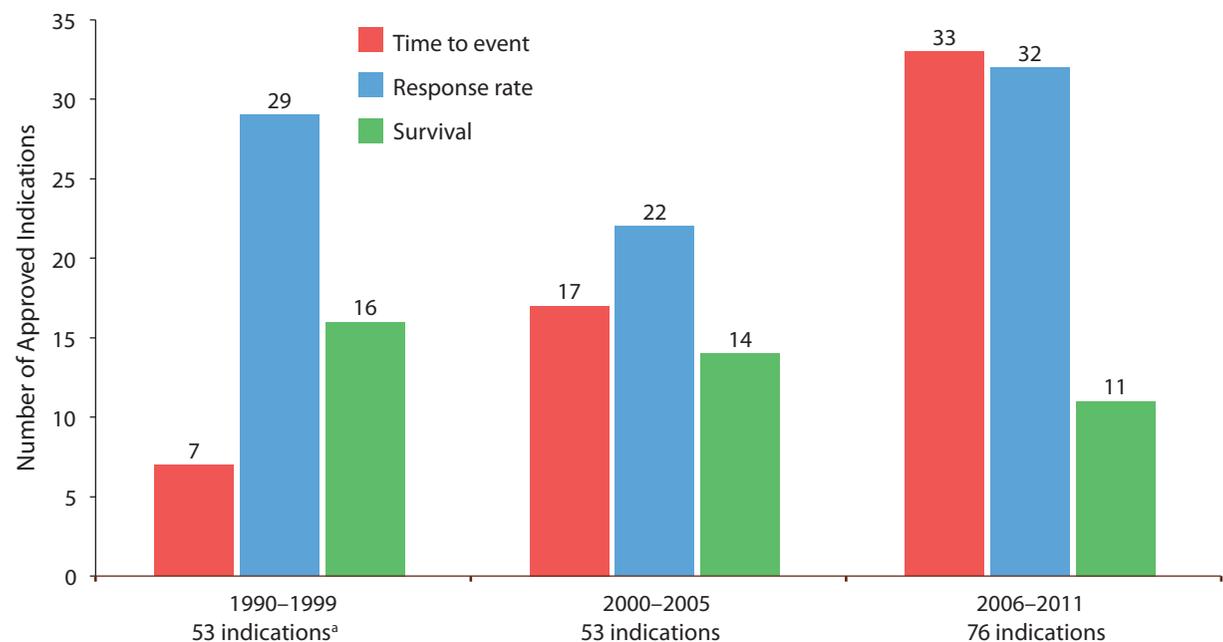
Endpoint Selection Process

The U.S. Food and Drug Administration (FDA) works in collaboration with pharmaceutical companies to select endpoints that are appropriate for the drug and the disease being

FIGURE 1

Change in the Use of Primary Endpoints for FDA-Approved Cancer Indications Over Time

Over the time span represented here, the percentage of trials with a time-to-event primary endpoint increased while the percentage of trials with survival as their primary endpoint decreased. The absolute number of trials with survival as their primary endpoint diminished slightly, however.



^aOne trial in this period used symptom palliation as a primary endpoint.

Adapted from Martell 2013

studied (McKee 2010), based on such criteria as the speed of disease progression and number of therapeutic options available. Furthermore, oncologists who help to design and run clinical trials may be consulted during this process. The agency also requires that the type of approval (regular vs accelerated) be considered in selecting the primary endpoint for a clinical trial (McKee 2010).

Surrogate endpoints have become widely used for regulatory approval of drugs (Figure 1). Accelerated approval has been used by the FDA since 1992, and regulations for accelerated approval allow for the use of surrogate endpoints (McKee 2010). The accelerated approval process includes the use of surrogate endpoints to allow therapies to reach patients sooner. The use of surrogate endpoints is not, however, limited to accelerated approval, and drugs that have gone through regular approval have also been approved based on surrogate endpoint outcomes.

Increasing Use of Surrogate Endpoints

For the 8-year period from 2005 through 2012, the FDA approved 41 indications for cancer drugs, accounting for 20% of all approvals during that period (Downing 2014). These approvals were supported by 55 pivotal trials and surrogate endpoints were used in 84% of them.

Extensive use of surrogate outcomes has continued in recent years (Figure 1). Excluding oncology drugs used to treat pain or other symptoms of cancer, 16 out of the 17 initial or additional indications for oncology drugs approved in 2014 through mid-2015 were based on a surrogate endpoint (FDA 2015):

- Progression-free survival (PFS): 6
- Event-free survival (EFS): 1
- Overall response rate (ORR): 8
- Complete remission with partial hematologic recovery rate: 1

The growing use of surrogate endpoints has prompted research into their relationship to survival. One recent study analyzed a sample of 65 studies by the strength of the correlation between surrogate markers and survival. In this example, results showed that in 48% (31 of 65) of the studies, the strength of the correlation was either medium or strong (Prasad 2015).

Surrogate Endpoints and Managed Care

The use of surrogate endpoints affects managed care organizations by increasing the number of drugs that may be considered for coverage. In addition, the growing variety of endpoints adds complexity to coverage decisions which were mostly based on OS. Furthermore, groups like the National Comprehensive Cancer Network® (NCCN®), which develop compendia accepted by the Centers for Medicare & Medicaid Services, may rely on research that includes surrogate endpoints as the rationale for their recommendations.

INDIVIDUAL ENDPOINTS

This section discusses individual endpoints, providing definitions, background, and some of their advantages and disadvantages. The individual endpoints are grouped into 3 categories: patient-centered, tumor-centered, and emerging tumor-centered.

PATIENT-CENTERED ENDPOINTS

Overall survival (OS). OS—widely regarded as the “gold standard” among primary endpoints—is defined as the time from randomization until death from any cause and is documented by the date of death (McKee 2010). OS can be measured in 2 ways: either as median OS, which is a duration of time at which 50% of patients in the trial are alive, or as a percentage of patients alive at different time points during the trial, which may be measured at 1, 2, or 5 years. Median OS is often used as a primary or co-primary endpoint. In some cases, post-marketing studies will continue in order to capture OS after initial efficacy is validated.

OS usually is measured in the intention-to-treat (ITT) population (McKee 2010). An advantage of using the ITT population to assess OS is that it captures patients who might withdraw from a trial because of treatment-related toxicity and whose deaths afterward might have been related to the effects of that toxicity (Gill 2006).

OS has several advantages as an endpoint. It is precise and easy to measure (McKee 2010). The outcomes are mutually exclusive (surviving vs not surviving). Moreover, investigator bias and subjective interpretation are not possible (Driscoll 2009).

But as an oncology endpoint, OS also has disadvantages (Fiteni 2014). When the prognosis of patients enrolled in a clinical trial is good, deaths are rare, so designing a clinical trial that will be adequately powered could require such a large number of patients that it would not be feasible to conduct. The fact that patients are surviving longer also means that the time required to see a treatment effect when OS is used as the endpoint may be significantly extended. In addition, because patients are treated with a wide variety of treatments both before and after participating in a trial, it is difficult to separate the effect of the investigational agent on OS from the effect that other treatments may have had.

Health-related quality of life (HR-QOL). HR-QOL endpoints measure physical and psychological status, participation in social activities, and other indicators of well-being, such as the ability to work. HR-QOL endpoints are important because many patients with cancer survive for years (Wilson 2015a). HR-QOL assessments may be used in a wide range of circumstances, from primary adjuvant treatment to palliative treatment of metastatic disease (Osoba 2011); however, instead of being primary outcomes in their own right, they tend to supplement biomedical outcomes by describing patients’ treatment experiences. Demonstrating improvement in HR-QOL takes on greater importance to stakeholders when OS benefit is small (Wilson 2015b).

Unlike most other endpoints, HR-QOL endpoints capture the effects of adverse events (AEs), albeit indirectly. It has been recommended that HR-QOL data collection include a period after the end of study treatment and during any subsequent treatment to account for disease-related symptoms requiring treatment after disease progression (Wilson 2015a). This extension may be important because a delay in disease progression might reduce the risk of some treatment-related AEs due to the reduced need for the treatment. Late-onset AEs could devalue initial HR-QOL improvement stemming from treatment.

One disadvantage of HR-QOL endpoints is that they may sometimes be difficult to compare across studies because a number of different instruments are used to assess HR-QOL (Wilson 2015b).

TUMOR-CENTERED ENDPOINTS

Progression-free survival (PFS) and time to progression (TTP). These endpoints are similar, but there are some impor-

tant differences. PFS is time from randomization to disease progression or death. TTP is time from randomization to time of disease progression, so PFS includes deaths, while TTP does not. Because PFS includes death and hence may correlate better with OS, the FDA may prefer PFS over TTP (Pazdur 2008).

Because of the possibility of the subjectivity of endpoint assessment, both of these endpoints should be assessed in randomized trials, preferably blinded, and patients must be evaluated on the same regular schedule and with the same techniques. Discussions with the FDA should focus on the magnitude of difference in TTP or PFS that would be considered clinically important because a statistically significant difference in TTP or PFS between treatment arms does not necessarily translate into a clinical benefit (Pazdur 2008).

As with OS, PFS can be measured either as median (the time at which 50% of the trial participants are alive or have not experienced disease progression) or at predetermined points of time (eg, the percentage of patients who are alive

Hazard Rates, Hazard Ratios, and Survival Curves

A survival analysis often is a key component of clinical trials in oncology. This analysis is concerned with the time from some starting point (eg, the initiation of treatment) to an event (eg, cure, hospitalization, or death). A common method of reporting survival is to plot the proportion of patients still alive over time in each study arm using a survival curve known as a Kaplan-Meier curve.

Sometimes survival is captured in a single number, such as the median survival, which is the time at which 50% of patients in a trial are alive. Another approach to calculating survival is with “landmark” measures of the number of people alive at a predetermined time.

But the Kaplan-Meier curve is a statistical picture of the percentage of patients surviving over a period of time; it cannot be summed up with a single number such as median survival or a landmark measure. Still, a measure of the entire curve can be estimated. The slope of the curve is the overall rate of death or risk of death; this is called the hazard ratio or the hazard rate.

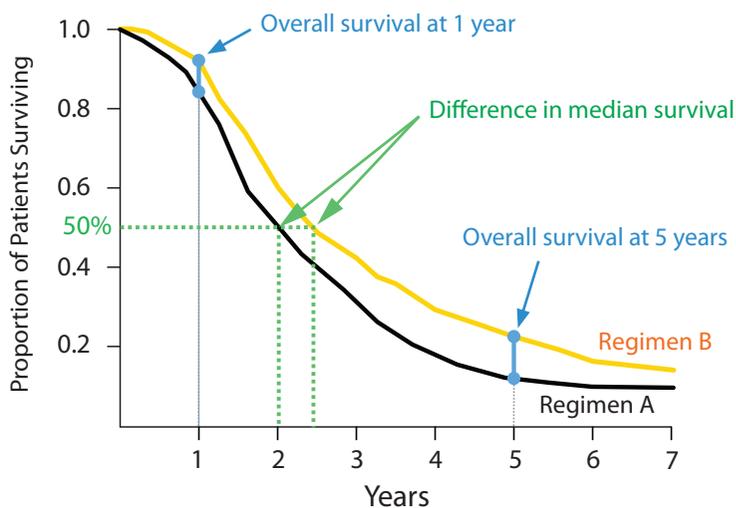
If a study has 2 arms, then 2 survival curves can be constructed, each with its own hazard rate. The hazard ratio (HR) is the ratio of the hazard for the

study drug divided by the hazard for the control. If HR=1, the hazard or risk of death in the 2 groups is equal. If HR >1, the risk of death is increased in the study group compared with the control group, while HR <1 means the risk of death is decreased in the study group compared with the control group.

The figure below provides an illustrative example of a Kaplan-Meier curve to show the difference between the median and landmark OS measurements.

The difference in median survival times appears to be relatively small, as does the difference in OS at 1 year (93% vs 85%), whereas at 5 years the OS difference is 23% vs 12%. Because of the shape of the curves, however, there also may be a large difference in OS as measured by the HR. It is important to look beyond single points along a survival curve and consider the HR, which analyzes the entirety of survival curves.

For illustrative purposes only



Commentary

Surrogate Endpoints From a Medical Director's Perspective

By Fredrick Allen May, MD

In this Clinical Brief, the breadth of current and future clinical endpoints has been well documented. The endpoints currently used most often in clinical trials fall into 2 basic groups, patient-centered and tumor-centered. The patient-centered endpoints have been used traditionally. But in today's environment of drug development, they can be problematic.

Using overall survival as an endpoint may not be practical in many cases because patients with cancer are living longer. As a result, statistically significant differences in survival among treatment groups may not appear in a reasonable amount of time. And while researchers are making progress toward measuring it objectively, quality of life is inherently subjective, shaped by values and experience—and no amount of measurement can change that.

Many of the tumor-centered endpoints are referred to as surrogate endpoints. They are often easier to achieve in smaller populations and in less time. As a result, they speed the development and approval of drugs so more treatments are available to patients with cancer who need effective therapy. However, the correlation between these endpoints and survival is not always high. Ask any patient or clinician who treats patients with

cancer. The goals of cancer therapy must be to extend life, give the patient as good a quality of life as possible, or both.

With that in mind, how do we balance that goal with the fact that 84% of recently approved cancer drugs were supported by a surrogate endpoint (Downing 2014)? How do we answer patients when the question of "outcomes" comes up in the discussion?

One possible strategy is to use postmarket studies that collect survival and other kinds of data. That way, promising drugs can get to market faster, but also be evaluated to make sure the surrogate endpoints are, in fact, reliable surrogates for what matters most—longer life and quality of life. We have seen cases when the FDA removed its approval after the surrogate endpoint turned out not to correlate well with survival and quality of life.

Understanding endpoints is difficult. In my experience, even oncologists sometimes have some trouble, especially when it comes to the newer ones. Patients are increasingly engaged in their care and asking questions. Providers and payers need to be prepared to explain endpoints and the

decisions upon which they are based.

So how does this relate to managed care? Health plans are pivotal gatekeepers in regulating the access to these exceptional drugs. We must understand the criteria—including the endpoints—that the FDA uses for approval. If we do, we can work knowledgeably with providers to get the most appropriate therapy to those who will benefit and avoid cases where the medication has little, if any, value.

This is no easy task. We must weigh efficacy, safety, and cost in a way that benefits our subscribers in the most efficient way possible.

Theodor Geisel, aka Dr. Seuss, put it well:

*So be sure when you step.
Step with care and great tact.
And remember that Life's
A Great Balancing Act.*

Dr. May is the medical director for Educators Mutual Insurance Association in Murray, Utah, and former medical director at Blue Cross & Blue Shield of Mississippi.

Dr. May received a stipend for this commentary.



Fredrick Allen May, MD

or have not experienced disease progression 1, 2, or 5 years after they were enrolled in the trial).

The advantages of PFS include the fact that it can be assessed before OS (FDA 2007), potentially allowing for shorter trials and bringing drugs to market sooner (Driscoll 2009). In addition, because more patients reach the endpoint, for any given patient population, the magnitude of the effect on PFS can be more easily seen than the effect on OS (FDA 2007). PFS seems to be most useful as a surrogate for OS when median survival after progression is relatively short, as in some advanced cancers (Wilson 2015b).

However, PFS as an endpoint has shortcomings. There's a subjective aspect to PFS, because it depends, in part, on when and how often patients are monitored with imaging

and other tests for disease (Driscoll 2009). Less frequent monitoring may identify progression later than monitoring that occurs more often. In addition, improvement in PFS may not correspond with improvement in OS (Booth 2012). The increase in tumor size measured at the time of progression may be so small that it has no bearing on how long the patient will live; the tumor burden may become lethal only long after PFS has been reached. Second, studies may be underpowered and incapable of detecting an absolute improvement in OS comparable to that found in PFS. If follow-up is long enough, a statistically significant hazard ratio for PFS should be compared with a similar hazard ratio for OS. Lack of this association in many trials suggests factors other than study power are involved (Booth 2012).

Disease-free survival (DFS). Definitions vary, but DFS is usually defined as time from randomization until recurrence or death from any cause (FDA 2007).

When DFS is used, it is usually in trials assessing adjuvant treatment after definitive surgery or radiotherapy (FDA 2007). DFS may be used as an endpoint in trials when long-term survival makes OS measurement impractical (FDA 2007). DFS may also be an acceptable and clinically relevant endpoint when a drug is being evaluated as adjuvant therapy for a condition in which it would take a long time to show an OS benefit (Wilson 2015b). DFS requires a randomized study; blinded review is recommended (FDA 2007).

Objective response rate (ORR) is sometimes called objective overall response rate. ORR measures the proportion of patients with a reduction in tumor size by a predefined amount (using standardized criteria, such as RECIST [Response Evaluation Criteria in Solid Tumors]) and for a minimum duration, usually measured from the time of treatment initiation to documented disease progression.

The FDA usually defines ORR as the sum of partial responses (PR) and complete responses (CR). This definition excludes stable disease (which is better assessed via TTP or DFS) and excludes minimal responses. These exclusions may enable ORR to be regarded as directly attributable to drug effect because stable disease may reflect the natural history of the disease rather than drug effect (McKee 2010).

ORR is often used in single-arm trials in refractory cancer (FDA 2007). However, because ORR is radiographically assessed, it may be unreliable because of the uncertainty of radiographic tumor assessment with some advanced cancers (McKee 2010).

Duration of response (DoR). DoR is the time from documentation of tumor response (either CR or PR) to disease progression (McKee 2010). When ORR is the primary endpoint in registration trials, the FDA may use response duration in its evaluation in addition to the percentages of complete and partial responses. DoR is one of the criteria the FDA may consider in determining adequacy of a surrogate endpoint for accelerated or regular approval.

DoR may be reported in conjunction with complete and/or partial responses that comprise ORR. Measurement of DoR is influenced by the frequency of follow-up after baseline evaluation, which is, in turn, affected by disease types and stages, treatment periodicity, and standard practice. The limitation of the precision created by variations in measured DoR should be taken into account if comparisons between trials are made (Eisenhauer 2009).

Time to treatment failure (TTF). TTF is a composite endpoint measuring time from randomization to treatment discontinuation for any reason (eg, disease progression, treatment toxicity, or death). The FDA no longer recommends TTF as an endpoint for gaining drug approval because it fails to clearly distinguish efficacy from toxicity, intolerance, and withdrawal from the study (FDA 2007).

EMERGING TUMOR-CENTERED ENDPOINTS

Immune-related response criteria (irRC). Tumor response criteria were first developed in the 1980s out of necessity to create a common language to describe the results of cancer treatment and provide a basis for advances in cancer therapy. Because of the promise of new immunotherapeutic agents in solid tumors, it is important to have clear and agreed-upon terms for immune-related tumor response assessment (Nishino 2013).

The irRC are endpoints intended to overcome the shortcomings of RECIST when applied to immunotherapies. They may be helpful in circumstances when if traditional response criteria are used to assess an immunotherapeutic agent, an unconventional response (eg, a temporary increase in tumor size from the arrival of immune cells at the tumor site) could lead to premature discontinuation of the trial. The irRC utilize an important concept: the overall tumor burden. The overall tumor burden embraces the combined size of index lesions present at baseline plus any new tumors detected after treatment begins (Hoos 2010). Under RECIST, these new tumors would be regarded as disease progression—indicating treatment failure—but irRC treats new tumors as part of the tumor burden instead of considering them as notification that the disease has worsened (Hoos 2012).

The irRC typically include 4 different kinds of response: immune-related complete response (irCR); immune-related partial response (irPR); immune-related stable disease (irSD); and immune-related progressive disease (irPD) (Hoos 2010).

Minimal residual disease (MRD). MRD is a new approach to detect traces of certain blood cancers using newer, highly sensitive technologies. Following completion of treatment, some patients may have persisting cancer cells that fall below the detection limits of standard laboratory methods but that newer technologies can detect. This remaining tumor burden is known as MRD. MRD technologies stratify patients by risk in order to improve outcomes and to reduce the risk of short- and long-term toxicity of therapy (FDA 2012).

The FDA and the American Society of Clinical Oncology sponsored a workshop in 2012 to discuss early MRD as a surrogate endpoint. According to the FDA, detection of MRD at an early time point “has emerged as a powerful and independent predictor of prolonged event-free survival (EFS)” (FDA 2012). EFS is the time between the end of primary treatment of a cancer and the onset of certain complications or events that the treatment was meant to prevent or delay.

MRD can be incorporated into a variety of endpoints, depending on the type of cancer and the technology used. One study used 4 endpoints based on MRD:

- Percentage of participants with an MRD response within 4 cycles of treatment (primary endpoint)
- Percentage of participants with an MRD response after each treatment cycle (secondary endpoint)

- Time to MRD progression (secondary endpoint)
- Time to MRD relapse (secondary endpoint)

MRD negativity may be an important criterion to evaluate treatment efficacy in hematologic tumors; it has been shown to correlate with survival in multiple clinical studies. The utility of MRD will continue to be evaluated as more data from clinical trials become available.

Central nervous system (CNS) endpoints. The blood-brain barrier prevents many therapeutic agents from reaching cancer cells in the brain (Deeken 2007), including metastases from various cancers (Schouten 2002). Moreover, the incidence of brain metastases is increasing as patients with cancer live longer (Deeken 2007). Although it once was common practice to exclude patients with brain metastases from clinical trials (Lin 2015), today researchers are conducting clinical trials to investigate drugs that might provide effective treatments for cancers that have metastasized to the brain. A common primary endpoint in these trials is the CNS ORR, using RECIST version 1.1, which sets the criteria for assessing tumors seen on imaging (NIH 2015). Another CNS primary endpoint is the CNS disease control rate (defined as the number of patients with stable disease, a partial response, or a complete response divided by the number of evaluable patients), also known as the CNS clinical benefit rate (NIH 2015). Both of these endpoints also are used as secondary endpoints, as are time to CNS progression as determined with RECIST 1.1, duration of response in the CNS, clinical benefit rate in the CNS, and PFS in the CNS, among others (NIH 2015).

Although CNS endpoints are important for the evaluation of drugs active in the CNS, they have their shortcomings. For example, use of intracranial PFS may not be a reliable surrogate for OS in patients with brain metastases (Lin 2013a). Intracranial PFS may be limited as an endpoint because it may not distinguish true progression from the effects of radiation treatment and may not capture the effect of treatment on extracranial progression (Lin 2013b). A drawback of RECIST when it is applied to primary brain tumors or brain metastases is that it doesn't consider steroid use or neurologic symptoms (Lin 2013b).

Pathological complete response (pCR). The pCR endpoint may be used to assess the efficacy of drugs given as neoadjuvant treatments.

The effectiveness of adjuvant therapy for certain cancers is well established, but certain subpopulations of patients continue to be at substantial risk of recurrence and death even with the best available adjuvant therapy. Unfortunately, novel postoperative systemic therapies for these patients can be assessed only in multiyear trials, and there is no early marker of potential efficacy in the adjuvant setting. In contrast, when systemic therapy is given in the preoperative setting, a pCR endpoint may be reasonably likely to predict clinical benefit in a high-risk population in a shorter time frame to support approval.

For example, the FDA has issued a guidance document about the use of pCR as a surrogate endpoint to support accelerated approval of drugs for neoadjuvant treatment of high-risk, early-stage breast cancer (FDA 2014). The FDA working group analyzed the relationship between pCR and long-term outcomes by pooling data for 13,000 patients enrolled in large trials of neoadjuvant therapy in which pCR was clearly defined and for which long-term follow-up data were available (FDA 2014). As a result of this analysis, the FDA has recognized 2 definitions of pCR for designing trials:

- Absence of residual invasive cancer on hematoxylin and eosin evaluation of the complete resected tissue specimen and all sampled regional lymph nodes following completion of neoadjuvant systemic therapy.
- Absence of residual invasive and in situ cancer on hematoxylin and eosin evaluation of the complete resected tissue specimen and all sampled regional lymph nodes following completion of neoadjuvant systemic therapy.

These definitions permit flexibility in terms of the approach to the surgical sites but reflect the fact that any residual invasive cancer following neoadjuvant therapy is indicative of a poorer prognosis (FDA 2014). However, rates of pCR measured in clinical trials may vary, depending on the FDA definition used to describe it.

DISCUSSION

This article has focused on the use of endpoints in clinical trials leading to regulatory approval of drugs used to treat cancer. In the future, there may be other uses for surrogate endpoints. They may play a larger role in early-stage trials or in highly stratified disease states to help identify appropriate trial populations. Although the demonstration of improvement in OS seen with some therapies with molecular targets may be difficult to achieve in many other cancers, appropriate selection of patient populations and surrogate endpoints could mean that large differences in outcomes will be apparent even in relatively small randomized trials (Wilson 2015a). To help identify these populations, surrogate endpoints could be employed in trials with novel designs that focus on the characteristics of subgroups—exceptional responders and nonresponders instead of the group as a whole—to gain a better understanding of response and resistance to therapy (Wilson 2015a).

Such an approach could lead to patient stratification in clinical trials according to type of genetic mutation instead of histology (Wilson 2015a). Toward this end, it has been proposed that surrogate endpoints such as PFS be incorporated in the first stage of multiarm, multistage trial designs to determine advancement to the next stage, where the outcome would be survival (Parmar 2008).

As medical science advances, some endpoints used in oncology trials may be used less often because they don't capture important effects of new treatments. Moreover, ad-

ditional endpoints may be identified as clinical trials continue to include novel therapies. Endpoints in development may look at other efficacy and safety measures, such as a CNS response. Some experts suggest periodic reviews of surrogate endpoints to ascertain that they are clinically relevant (Wilson 2015a). Meanwhile, new surrogate endpoints are being introduced. Payers and physicians will need to become familiar with them to evaluate new drugs in a meaningful way when trials using new endpoints are used for regulatory approval.

CONCLUSION

Surrogate endpoints have become an important tool for the development of oncology drugs and their regulatory approval. Although OS is considered the “gold standard” endpoint, it has become more difficult to use and measure,

partly because survival is influenced by the growing number of therapies available to patients before and after a trial. OS may also be affected by patients in control groups crossing over to new therapies after trial results have shown them to be effective. In addition, with increased patient survival, trials may have to be larger, longer, or both, to capture enough survival data to reach a statistically significant result. For these reasons, surrogate endpoints are now used in more oncology clinical trials.

Surrogate endpoints may shorten trials and time to approval. However, shortening the time it takes a new drug to reach patients is important only if it provides true clinical benefit (Wilson 2015a). As new treatments and technologies emerge, pharmaceutical companies and regulators are working to make sure that surrogate endpoints are clinically relevant.

TABLE 1
Important Cancer Approval Endpoints

Endpoint	Regulatory evidence	Study design	Advantages	Disadvantages
Overall survival (OS)	Clinical benefit for regular approval	<ul style="list-style-type: none"> • Randomized studies essential • Blinding not essential 	<ul style="list-style-type: none"> • Universally accepted direct measure of benefit • Easily measured • Precisely measured 	<ul style="list-style-type: none"> • May involve larger studies • May be affected by crossover therapy and sequential therapy • Includes noncancer deaths
Symptom endpoints (patient-reported outcomes)	Clinical benefit for regular approval	<ul style="list-style-type: none"> • Randomized blinded studies 	<ul style="list-style-type: none"> • Patient perspective of direct clinical benefit 	<ul style="list-style-type: none"> • Blinding is often difficult • Data are frequently missing or incomplete • Clinical significance of small changes is unknown • Multiple analyses • Lack of validated instruments
Disease-free survival (DFS)	Surrogate for accelerated approval or regular approval ^a	<ul style="list-style-type: none"> • Randomized studies essential • Blinding preferred • Blinded review recommended 	<ul style="list-style-type: none"> • Smaller sample size and shorter follow-up necessary compared with survival studies 	<ul style="list-style-type: none"> • 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
Objective response rate (ORR)	Surrogate for accelerated approval or regular approval ^a	<ul style="list-style-type: none"> • Single-arm or randomized studies can be used • Blinding preferred in comparative studies • Blinded review recommended 	<ul style="list-style-type: none"> • Can be assessed in single-arm studies • Assessed earlier and in smaller studies compared with survival studies • Effect attributable to drug, not natural history 	<ul style="list-style-type: none"> • Not a direct measure of benefit • Not a comprehensive measure of drug activity • Only a subset of patients who benefit
Complete response (CR)	Surrogate for accelerated approval or regular approval ^a	<ul style="list-style-type: none"> • Single-arm or randomized studies can be used • Blinding preferred in comparative studies • Blinded review recommended 	<ul style="list-style-type: none"> • Can be assessed in single-arm studies • Durable complete responses can represent clinical benefit • Assessed earlier and in smaller studies compared with survival studies 	<ul style="list-style-type: none"> • Not a direct measure of benefit in all cases • Not a comprehensive measure of drug activity • Small subset of patients with benefit
Progression-free survival (PFS) (includes all deaths) or Time to progression (TTP) (deaths before progression censored)	Surrogate for accelerated approval or regular approval ^a	<ul style="list-style-type: none"> • Randomized studies essential • Blinding preferred • Blinded review recommended 	<ul style="list-style-type: none"> • 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 	<ul style="list-style-type: none"> • 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
Emerging endpoints				
Immune-related response criteria (irRC)	Surrogate for accelerated approval, usually as secondary endpoint	<ul style="list-style-type: none"> • Total tumor burden should be measured at many time points, not just a single point 	<ul style="list-style-type: none"> • Captures data impossible to measure with other endpoints • Appropriate for cytostatic agents that produce unconventional responses 	<ul style="list-style-type: none"> • Frequent assessments
Minimal residual disease (MRD)	Surrogate for accelerated approval, usually as secondary endpoint	<ul style="list-style-type: none"> • Single-arm or randomized studies can be used • Samples should be taken at different time points in the study 	<ul style="list-style-type: none"> • Identifies traces of cancer that elude other tests • Facilitates stratification of subjects in clinical trials 	<ul style="list-style-type: none"> • Requires advanced technologies
Central nervous system (CNS)	Surrogate for accelerated approval, usually as secondary endpoint		<ul style="list-style-type: none"> • Provides a way to assess drugs designed to treat brain metastases 	<ul style="list-style-type: none"> • Intracranial PFS may not be reliable surrogate for OS • Standard RECIST criteria may identify pseudoprogression instead of actual progression
Pathological complete response (pCR)	Surrogate for accelerated approval, usually as secondary endpoint	<ul style="list-style-type: none"> • Randomized controlled trial intended to show superiority (neoadjuvant breast cancer setting) 	<ul style="list-style-type: none"> • Can be assessed quickly, sometimes in months instead of years 	<ul style="list-style-type: none"> • Generally based on objective and quantitative assessment • Definitions vary among studies • Not predictive of outcomes in most breast cancers

^aAdequacy 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.

Sources: Cortazar 2014, FDA 2007, FDA 2014, Hoos 2010, Hoos 2012, Klempner 2015, Lin 2013a, Lin 2013b

References

- Booth CM, Eisenhauer EA. Progression-free survival: meaningful or simply measurable? *J Clin Oncol*. 2012;30(10):1030–1033.
- Cortazar P, Zhang L, Untch M, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet*. 2014;384:164–172.
- Deeken JF, Löscher W. The blood-brain barrier and cancer: transporters, treatment, and Trojan horses. *Clin Cancer Res*. 2007;13(6):1663–1674.
- Downing NS, Aminawung JA, Shah ND, et al. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005–2012. *JAMA*. 2014;311(4):368–377.
- Driscoll JJ, Rixe O. Overall survival: still the gold standard: why overall survival remains the definitive end point in cancer clinical trials. *Cancer J*. 2009;15(5):401–405.
- Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumors: Revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45(2):228–247.
- FDA (US Food and Drug Administration). Drugs@FDA. FDA approved drug products. www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm. Accessed September 28, 2015.
- FDA. Minimal Residual Disease (MRD) as a Surrogate Endpoint in Acute Lymphoblastic Leukemia (ALL) Workshop. April 18, 2012. <http://www.fda.gov/downloads/Drugs/NewsEvents/UCM299917.pdf>. Accessed September, 22, 2015.
- FDA. 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 September, 22, 2015.
- FDA. Guidance for Industry. Pathological Complete Response in Neoadjuvant Treatment of High-Risk Early-Stage Breast Cancer: Use as an Endpoint to Support Accelerated Approval. Center for Drug Evaluation and Research. October 2014. <http://www.fda.gov/downloads/DrugsGuidanceComplianceRegulatoryInformation/Guidances/UCM305501.pdf>. Accessed September 22, 2015.
- Fiteni F, Westeel V, Pivot X, et al. Endpoints in cancer clinical trials. *J Visc Surg*. 2014;151(1):17–22.
- 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.
- Hoos A. Evolution of end points for cancer immunotherapy trials. *Ann Oncol*. 2012;23(suppl 8):viii47–viii52.
- Hoos A, Eggermont AM, Janetzki S, et al. Improved endpoints for cancer immunotherapy trials. *J Natl Cancer Inst*. 2010;102(18):1388–1397.
- Klempner SJ, Ou SH. Anaplastic lymphoma kinase inhibitors in brain metastases from ALK+ non-small cell lung cancer: hitting the target even in the CNS. *Chin Clin Oncol*. 2015;4(2):20.
- Lin NU, Lee EQ, Aoyama H, et al; Response Assessment in Neuro-Oncology (RANO) group. Response assessment criteria for brain metastases: proposal from the RANO group. *Lancet Oncol*. 2015;16(6):e270–e278.
- Lin NU, Lee EQ, Aoyama H, et al; Response Assessment in Neuro-Oncology (RANO) group. Challenges relating to solid tumour brain metastases in clinical trials, part 1: patient population, response, and progression. A report from the RANO group. *Lancet Oncol*. 2013b;14(10):e396–e406.
- Lin NU, Wefel JS, Lee EQ, et al; Response Assessment in Neuro-Oncology (RANO) group. Challenges relating to solid tumour brain metastases in clinical trials, part 2: neurocognitive, neurological, and quality-of-life outcomes. A report from the RANO group. *Lancet Oncol*. 2013a;14(10):e407–e416.
- Martell RE, Sermer D, Getz K, Kaitin KI. Oncology drug development and approval of systemic anticancer therapy by the U.S. Food and Drug Administration. *Oncologist*. 2013;18(1):104–111.
- McKee AE, Farrell AT, Pazdur R, Woodcock J. The role of the U.S. Food and Drug Administration review process: clinical trial endpoints in oncology. *Oncologist*. 2010;15(suppl 1):13–18.
- NIH (U.S. National Institutes of Health). ClinicalTrials.gov. <https://clinicaltrials.gov>. Accessed October 12, 2015.
- Nishino M, Giobbie-Hurder A, Gargano M, et al. Developing a common language for tumor response to immunotherapy: immune-related response criteria using unidimensional measurements. *Clin Cancer Res*. 2013;19(14):3936–3943.
- Osoba D. Health-related quality of life and cancer clinical trials. *Ther Adv Med Oncol*. 2011;3(2):57–71.
- Parmar MK, Barthel FM, Sydes M, et al. Speeding up the evaluation of new agents in cancer. *J Natl Cancer Inst*. 2008;100:1204–1214.
- Pazdur R. Endpoints for assessing drug activity in clinical trials. *Oncologist*. 2008;13(suppl 2):19–21.
- Prasad V, Kim C, Burotto M, Vandross A. The strength of association between surrogate end points and survival in oncology: a systematic review of trial-level meta-analyses. *JAMA Intern Med*. 2015;175(8):1389–1398.
- Schouten LJ, Rutten J, Huvencers HA, Twijnstra A. Incidence of brain metastases in a cohort of patients with carcinoma of the breast, colon, kidney, and lung and melanoma. *Cancer*. 2002;94(10):2698–2705.
- Wilson MK, Collyar D, Chingos DT, et al. Outcomes and endpoints in cancer trials: bridging the divide. *Lancet Oncol*. 2015a;16(1):e43–e52.
- Wilson MK, Karakasis K, Oza AM. Outcomes and endpoints in trials of cancer treatment: the past, present, and future. *Lancet Oncol*. 2015b;16(1):e32–e42.

