Trends-in-Medicine

December 2003
By Lynne Peterson

SUMMARY

The ODAC panel reviewed the findings of an April 2003 meeting on lung cancer endpoints and spent a full day discussing issues the FDA wanted clarified. Much of the discussion focused on time-to-progression (TTP) as an endpoint, and the panel decided that progression free survival (PFS) is a better endpoint than TTP. A panel felt PFS could support both regular and accelerated drug approval in metastatic NSCLC but not inoperable NSCLC. The panel also recommended that disease free survival (DFS) be permitted as an endpoint for regular drug approval.

Trends-in-Medicine
Stephen Snyder, Publisher
1879 Avenida Dracaena
Jensen Beach, FL 34957
772-334-7409 Fax 772-334-0856
www.trends-in-medicine.com

CLINICAL TRIAL DESIGNS AND ENDPOINTS: FDA ONCOLOGIC DRUGS ADVISORY COMMITTEE (ODAC)

December 16, 2003
Rockville, MD

This was the third in a series of meetings by the FDA in discussing endpoints for clinical trials in oncology. The ODAC panel reviewed the findings of an April 2003 meeting on lung cancer endpoints and spent a full day discussing issues the FDA wanted clarified. Much of the discussion focused on time-to-progression (TTP) as an endpoint, and the panel decided that progression free survival (PFS) is a better endpoint than TTP. Details of the discussion follow and may be instructive in interpreting the results of future clinical trials.

The panel also took these votes:

1. For approval of drugs for first-line treatment of advanced NSCLC, could a progression-free survival (PFS) benefit of a new drug, compared to a standard first-line regimen, justify regular drug approval? (Assuming the standard control arm has a known small – 2 month – survival benefit)
   in metastatic NSCLC: **PASSED. Vote: Yes 10, No 8**
   in inoperable NSCLC: **FAILED. Vote: No 15, Yes 3**

2. If an improvement in PFS would not support regular approval, could it support accelerated approval?
PASSED. Vote: Yes 17, No 0, Abstain 1

3. In NSCLC, should a disease free survival (DFS) improvement from adjuvant chemotherapy support regular drug approval?
PASSED. Vote: Yes 17, No 1

THE FDA PERSPECTIVE

Dr. Richard Pazdur, Director of Oncology Drug Products for the FDA, was ill and participated by telephone from home. Therefore, Dr. Grant Williams, Deputy Director for the FDA's Division of Oncology Drug Products, Office of Drug Evaluation and Research I, opened the meeting, commenting, "We are excited about getting endpoints out and discussed. For us it is a difficult process...Trying to be consistent with the endpoints we require across the many settings is quite a challenge."

Approval mechanisms are:

1. Regular approval – a sponsor must show a clinical benefit or an established surrogate for clinical benefit. Usually that means improvement in quality or quantity of life (survival).

2. Accelerated approval – a surrogate that is reasonably likely to predict a clinical benefit in a serious and life-threatening illness. The new therapy must provide an advantage over available therapy. Post-marketing studies are required to verify the clinical benefit, and substantial evidence from well-controlled clinical trials are required – *not borderline evidence regarding a clinical benefit endpoint.*

The usual requirement is for two trials, but a single trial may sometimes suffice if there is:

- A large, multicenter trial
- Consistent results across study subsets
- Multiple studies in a single trial (e.g., factorial design)
- Results from secondary endpoints *also* are positive
- Statistically persuasive results

Single trials are acceptable for oncology supplemental applications for:

- Different stages of disease (e.g., metastatic vs. adjuvant settings)
- Different treatment settings (e.g., refractory vs. first-line therapy)
- Combination therapy vs. monotherapy
- Closely-related cancers

From 1990-2002, 73% of approvals were based on survival. There was one drug approved based on time-to-progression (TTP), and two approved on disease free survival (DFS). Endpoints for past approvals were:

- Survival has been the gold standard and provides an unambiguous endpoint that is easily measured
- DFS (adjuvant setting)
- Tumor-related symptoms/patient reported outcomes – are quite relevant from the patients' perspective. For example for mitoxantrone for HRPC (pain scale), Photofrin for obstructive lesions (dysphagia scale) and bisphosphenates (skeletal-related events)
- TTP (advantages and challenges)
- ORR (e.g., durable complete responses in leukemias and some solid tumors, partial response in hormonal therapy of breast cancer)

Most accelerated approvals have been based on objective response rate (ORR) in studies without an active comparator -- single arm studies -- but some randomized trials have been done [e.g., anastrozole (AstraZeneca's Arimidex) and oxaliplatin (Sanofi's Eloxatin)]. Two strategies have emerged for accelerated approval:

A. **Approval based on response rate in single arm study of refractory patients and a confirmatory study in a related population** (e.g., less refractory patients). This allows rapid completion of single arm studies. But this has become more and more difficult to evaluate marginal benefits. And evaluation in refractory populations first may cause us to miss an active drug. And there are other evaluation limitations.

B. **Evaluation of surrogate endpoints in a randomized study with clinical benefit established by final analysis of the same study.** This facilitates completion of the confirmatory study and allows comparison to available therapy and evaluation of the toxicity profile – but it may require more time and patients than a single arm study and accelerated approval could influence completion of the study.

Clinical trial issues include:

1. **Use of placebos.** A placebo-alone arm is usually not feasible in advanced cancer, but it potentially could be used for:
   a. Prevention, adjuvant or early disease for (b) add-on designs and (c) continuation of drug and placebo after failure of Drug A.
   b. No blind or placebo has consequences: Control must be an active drug, with a superiority design preferred. Non-inferiority requires large trials, and the quality of historical data limits design, so it is difficult to approve drugs that are similar but less toxic.

2. **Combination drug problem.**

3. **Non-inferiority.** Dr. Williams said, "We love superiority, and we love superiority trials. Equivalence is a word you should never say to a statistician. Equivalence is something that can never be proven. An important point is that proof of non-inferiority does not necessarily prove efficacy. A common problem in oncology journals is that 'no statistical difference' is considered the same as non-inferiority, and it isn't. Non-inferiority studies are not the FDA's favorite trial design. Violating the constancy assumption could lead to approval of a 'toxic placebo.' Sloppiness obscures differences and in non-inferiority, it could lead to a false efficacy claim…Determining the margin from historical cancer drug effects is difficult and leads to a very small margin and very large non-inferiority trials."

4. **Surrogate endpoints.** There are few validated surrogates in oncology.

There are numerous issues with the use of TTP. Dr. Williams said, "The question we should ask is not if improvement in TTP has clinical meaning -- no one in oncology disputes that delaying progression is a good thing – but:

- Can you reliably measure it?
- If you can measure it, what does it mean?
- How much progression delay is worth how much toxicity?

- What is the relative meaning of a TTP benefit to other benefits such as survival?"

TTP would be more persuasive if:

- Symptoms frequently occurred at or soon after progression time
- TTP increment is large
- Treatment toxicity is low
- The benefit of available drugs is less

Four years ago, ODAC felt TTP was not appropriate for regular approval but was appropriate for accelerated approvals. At the time, the panel raised questions about:

- Small TTP benefits with current drugs.
- Poor correlation with survival.
- Unreliable TTP measurements.
- Necessity for frequent measurements to achieve reliability.

Other points that were made relating to TTP:

- If TTP is to be used, there should be careful agreement between the sponsor and the FDA on the protocol, the statistical analysis plan, and the case report forms.
- If non-tumor deaths are not counted as events and are censored at the last visit, it makes an assumption that there is no relationship between death and progression, and that assumption might be questionable.

Pros and Cons of TTP

Positives	Negatives
Measured in all patients, so it might be a better measure of overall benefit than response	May not correlate with survival
Measures cytostatic activity	Indirect measure of patient benefit
Often the reason oncologists change therapy	Unclear meaning of a small difference
Assessed before crossover, which is growing in importance as we develop more effective drugs	Concerns with reliability in an unblinded setting (bias)
Requires smaller studies	Expensive to measure and difficult to verify
May have face validity	

Survival vs. TTP As Endpoint

Survival	TTP
100% accurate event	Less accurate
100% accurate time	Less accurate
Assed daily	Assessed every 2-6 months
Importance questioned	Uncertain
Both safety and efficacy	Only efficacy
Takes longer	Faster
Might be obscured by secondary therapy	Not obscured by secondary therapy

- The problem with measuring TTP without deaths -- and measuring deaths in a survival analysis – is that patients lost to follow-up cause prolongation of progression time, and careful follow-up is needed.
- Time-to-treatment-failure (TTF) is a composite endpoint, is not an acceptable endpoint for documenting efficacy, and does not support drug approval.

THE PANEL DEBATE

In a non-inferiority trial, does a sponsor have to prove less toxicity?

Dr. Williams said, "For us, it means you met the margin – that the drug works. It is a separate judgment as to whether you are less toxic. There is no direct requirement to be less toxic…You could have toxicity affect your margin; you might be willing to accept less efficacy if you knew something was less toxic." Dr. Pazdur added, "A lot of people confuse less toxic and non-inferior. The toxicity evaluation is different, and many times we see drugs that are not less toxic but have different toxicity." Dr. Temple added, "The grim reality for non-inferiority is that we want to preserve half the effect of what a drug is…It is a tremendous problem to get less toxicity or more easily taken drugs."

What happens if a drug has equivalent TTP but a different duration of response?

Dr. Williams said, "If it is the same TTP, that is one thing…but we've never had TTP as a primary endpoint and seen differences in response rate. Obviously, duration of response is always a big consideration."

An industry representative on the panel argued that TTP is a better endpoint for regular approval than survival, "Survival is plagued by a number of biases…One way to address the TTP/response issue is to analyze TTP for responders. When you look at TTP for responders, that is an even better endpoint than duration of response. TTP at least has a definite calendar date for onset."

Dr. Temple responded, "Certainly a long response is beneficial in certain cases – leukemia, testicular cancer, etc. -- but as an endpoint in clinical trials, we have never been successful in incorporating that measurement into our overall evaluation." However, FDA officials noted that duration of response was shown with IL-2, fludarabine (Berlex's Fludara) and Millennium's Velcade (bortezomib).

FDA Discussion Questions

What is the role of survival as a primary endpoint?
The panel agreed: Survival is the goal but not necessarily a good endpoint because it can be biased. If the original protocol is survival, then the confounding factors have to be ignored. Different guidelines need to be applied for each biological subset. At this point, the panel cannot demand a sponsor show a survival benefit.

Comments included:

- *The patient representative:* "Survival has more biases in my mind than TTP."
- *An oncologist:* "A meaningful increase in survival depends on which cancer you are talking about. Two months may be very meaningful in melanoma, but that's nothing in follicular lymphomas…You also have to consider whether it is front-line or relapse therapy. Four to six month survival in relapsed follicular lymphoma would be okay, but not in up-front follicular lymphoma…This is a totally moving target, especially in the hematologic malignancies…Every time a new drug is approved, the bar is set higher and higher…The 10% response rate with Iressa (AstraZeneca, gefitinib) would not cut it at all in hematologic malignancies."
- *Dr. Williams:* "We have not yet not approved a drug with a survival effect that we really believed – after trading off toxicity."
- *Statistician:* "The choice of endpoint ought to be something patients really care about. Cancer has a huge effect on duration of survival, and prolonging survival is of tremendous importance, but there may well be other measures…It may be that we could affect mechanisms that don't impact the clinical endpoints of interest…So, the primary endpoints for registration should reflect a tangible benefit or biologic activity that has been validated…The argument against survival is it may be impacted by subsequent interventions."
- *Dr. Temple:* "I'm very worried about survival where crossover is predictable. You would probably need huge studies…If survival is the endpoint, then everyone has to sit down and try to prevent crossovers or do large studies."

Do clinical settings exist where TTP improvement should be considered an established surrogate for clinical benefit and should support regular drug approval?
The panel felt PFS is a better endpoint than TTP. TTP requires a rigorous assessment and probably repeated assessments and can't be used for patients who are very symptomatic. If TTP is used, toxicity needs to be factored into the risk:benefit equation. In a disease with a low CR rate, the panel thought therapies unlikely to alter survival could be a primary endpoint but not in cases where standard therapy already has a benefit.

Comments included:

- *Dr. Temple:* "I'm not sure I agree that we don't expect these drugs to alter survival. It may be difficult to prove, but my assumption is if something has an effect on TTP, it probably does have a favorable effect on survival -- even if you can't measure that well…Sponsors using TTP, should be sure they are seeing people at a regular interval – every two or three months, etc…We have been encouraging people to look at time to symptomatic progression, and we have met with total failure…I don't know why. Symptomatic improvement has always been a valid endpoint, but, except for pain, prostate cancer and esophageal obstruction, we have had little success with this (endpoint)."
- *Chairperson:* "TTP is excellent in metastatic prostate cancer in elderly patients where no matter what you do, they die of non-cancer reasons and if you can keep them symptom-free, it is very valuable."
- *Patient advocate:* "If a patient starts highly symptomatic, and the treatment relieves symptoms and delays the time to them getting worse, then TTP has value to the patient…If a patient has profoundly bad symptoms and knows he could get worse but then doesn't, that is acceptable…I don't think you can have a better endpoint than PFS…Survival is not a good endpoint."
- *Oncologist:* "I would want a clinical benefit beyond TTP – symptom relief, less toxicity, better quality of life, etc. If it is just TTP without the other things, I'm not sure it is valid in a clinical sense."
- *Statistician:* "There is a missing data issue with TTP…I find TTP especially problematic for registration rather than as a supportive measure, so I would prefer PFS to TTP.

The FDA offered a series of possible scenarios for use of TTP, and the panel's statistician suggested the following would be appropriate:

- When many patients are symptomatic at time of progression (e.g., patients with bone metastases from prostate cancer).
- When the estimate of TTP benefit is large and precisely defined.
- When a new drug shows superior TTP to a standard drug.
- When progression time is determined in a blinded study.
- In an unblinded study, when progression time is determined by a blinded independent review group.
- When drugs have minimal toxicity (e.g., hormone therapies for breast cancer.

Is DFS generally an adequate endpoint for approval of cancer drugs or is additional evidence needed, such a data demonstration (or suggestion) that DFS is a survival surrogate? Is DFS an endpoint or a surrogate marker?
The panel agreed that DFS is an actual, not a surrogate endpoint, especially in leukemia patients. However, several panel members pointed out that DFS does not necessarily mean symptom improvement.

Does the adequacy of DFS vary with the clinical setting. For instance, where:

a. **No standard adjuvant therapy exists and treatment with the investigational drug shows superior DFS compared to an unproven control regimen. Yes.**
b. **Treatment with an investigational drug shows prolongation of DFS compared to highly effective standard therapy (that imparts a survival benefit). Yes.**
c. **Treatment with an investigational drug shows non-inferior DFS compared to highly effective standard therapy (that imparts a survival benefit). Mixed opinions.**

The panel felt that DFS can be a primary, not just a surrogate, endpoint. The chair concluded, "In the right circumstances, when DFS is the endpoint, we would not be so concerned with survival, but we want to look at survival data where it is available…DFS is best used where high response rates are expected, especially in people off therapy with high toxicity, and functionality is critical in looking at DFS…In randomized trials where the comparator is a highly effective therapy with a curative fraction, opinions are mixed opinion on whether DFS is an adequate EP or a surrogate marker."

Comments included:

- *Oncologist*: "It doesn't necessarily have to be better to be approvable."
- *Chair*: "It varies by disease, but generally, yes."
- *Oncologist*: "It depends on the disease…I can't give a blanket answer."
- *Another oncologist:* It seems to me that if there is a curative regimen at some level, and a new drug shows prolongation of DFS, in that setting I'd say DFS is a surrogate, and that is what accelerated approval is all about – something that is highly likely to convert to a survival benefit in the future, but is still a good surrogate marker…It seems to me functionality is the critical issue."
- *Dr. Temple:* "Durable CR is a recognized benefit, and we don't see that very often…but where it occurs, that is persuasive…All treatments for testicular cancer were approved on data like that."

THE ASCO PERSEPCTIVE

In addition to the open meeting in April 2003, the FDA and ASCO have had a series of telephone conferences. The FDA's Dr. Pazdur said, "Not all cancers are the same…In the future it is highly likely that we will have to look at endpoints in individual cancers based on data in those individual cancers…The (FDA's) oncology division has determined that new drugs using accelerated approval should show an advantage over existing drugs."

Dr. Paul Bunn, a lung cancer specialist and Immediate Past President of the American Society of Clinical Oncology (ASCO), reviewed those discussions. He said, "In my opinion, overall response (OR) can be considered in accelerated approvals because they likely indicate a patient benefit, but improved OR rates do not always translate to improved survival.

OR could be used if:

- In the first line setting, where active agents have OR rates of >20% in limited institution trials and >15% in multi-institution trials.
- In second line setting, where active agents have OR rates of >10% in limited and >8T in multi studies.
- The trial were large enough. To demonstrate an OR of 25% with 95% CI (+- 5%) would require 400-patient trials in a first-line, limited institution setting. A multi-institutional setting would require 625 patients.

Some of the differences between lung cancer and other cancers are that lung cancer patients generally:

- Present with advanced stage (IIIb or IV) disease.
- Are symptomatic.
- Have co-morbid disease (cardiopulmonary).
- Are elderly (>68 years old).
- Are difficult to recruit into surgical trials.
- Show an OR in <25% patients.
- Have not been influenced by second-line therapy until recently.

PATIENT-REPORTED OUTCOMES

These were discussed extensively at the April 2003 lung cancer endpoints workshop, and an expert reviewed that discussion. He said, "Patient reported outcomes (PROs) are ready for prime time…but we need to use care in the selection of the PRO instrument, and before the trial begins, you need to clearly delineate the primary and secondary endpoints and the analysis plan to be followed."

Drug	Limited institution Phase II single agent therapy					Multi-institutional Phase III trials				
	# of trials	# of patients	OR	MS	1 year survival	# of trials	# of patients	OR	MS	1 year survival
Docetaxel	8	300	26%	9.7%	22%	2	285	18%	7.0%	33%
Paclitaxel	9	370	27%	7.9%	3.8%	3	377	17%	6.3%	35%
Gemcitabine	12	572	21	9.7%	29%	6	736	16.3%	7.2%	28%
Vinorelbine*	14	621	20%	7.6%	24%	NA	769	18	7.2%	25%

*** The only single agent approved for NSCLC.**

Quality of life instruments for lung cancer:

- **Lung Cancer Symptom Scale** (LCSS). This was specifically developed for use in clinical trials. It has nine patient items and six observer items.
- **EORTC**. The general and lung cancer modules have 30-40 items and were developed for general use.
- **FACT-L**. Again, the general and lung cancer modules have 30-40 items and were developed for general use.
- **Rotterdam Symptom Checklist** (RSCL). It is not lung-cancer specific.
- **Hospital Anxiety and Depression Scale** (HADS). This is often used with RSCL in British Medical Research Council studies, but it is not lung-cancer specific.

Problems in evaluation of PROs in clinical trials include:

- Cumbersome instruments
- Patient deterioration
- Lack of investigator commitment

The Industry Perspective

An official with AstraZeneca made several points, including:

- Symptomatic endpoints should be allowed for approval when well-validated scales are used.
- Trials should be allowed in subsets of patients with performance status II. Inclusion and exclusion criteria currently eliminate these patients because of their short life expectancy, and many are unsuitable for cytotoxic therapy.
- ODAC's recommendation that PFS could be the sole basis for approval in certain situations is encouraging.

An efficacy standard for non-inferiority trials would increase trial size tremendously. He said, "We don't believe the answer is to avoid non-inferiority trials…We think there are areas where it would be the design of choice."

Miscellaneous

AstraZeneca's Iressa (gefitinib)

Several panel members were asked, in retrospect, how they feel about the approval of Iressa, and the responses were mixed. One panel member said he still isn't sure, but it looks as if the only patients for whom Iressa is useful are young non-smokers. Another thought approval was a bad idea. A third was uncertain.

Genentech/OSI's Tarceva (erlotinib)

Dr. Bunn, a past president of ASCO, argued that Tarceva should be approved, 'Erlotinib will come before this committee in a trial where the hazard rate for the study was a >30% reduction for a single pill in second or third line NSCLC...That may not make it against best supportive care in terms of survival, but I'll eat my hat if, in terms of response, it isn't statistically significant…If erlotinib has 9% response rate, and best supportive care has a 2% rate, I'd say accelerated approval should be given."

What if a drug like Tarceva (which was in mind but not mentioned) missed its primary endpoint of survival but showed a benefit on secondary endpoints of objective response and/or quality of life. A senior FDA official responded, "If the primary endpoint is missed, it is very difficult to get approved on secondary endpoints because the possibility of a false answer is magnified. If there is a trend with the primary endpoint, and the company 'blew our socks off' with a secondary endpoint, it might be approvable." The chairman of the ODAC panel said, "The secondary endpoints would have to be compelling, and I don't just mean statistically significant. They would have to have a wide margin of benefit."

What kind of message will it send to sponsors if Tarceva misses its primary survival endpoint? A pharma official predicted it would put a chill on cytostatic agents and cause sponsors to concentrate on cytostatics.

Genta's Genasense and Genentech's Avastin (bevacizumab)

Taking a drug directly from Phase I to Phase III was criticized – unless the drug (like Genasense) showed compelling pre-clinical results. However, this may change. ASCO's Dr. Bunn said, "Bevacizumab may be the first to prove this wrong." ♦