Trends-in-Medicine

April 2003
By Lynne Peterson

SUMMARY

Time-to-progression (TTP) will not be adopted as an endpoint for lung cancer trials in the near future. Quality of life and patient-report-outcomes are good supporting endpoints, but they also are not ready for use as primary endpoints. Non-inferiority trials, particularly the increasingly common single-arm trials, are difficult to interpret, and the FDA wants strong data, but no specific p-value has been set. Accelerated approvals are likely to be harder to get; ODAC will be more critical of data in these cases, and companies generally will have to have confirmatory studies underway before an accelerated approval is granted. This meeting was ***not*** a prelude to either approval or non-approval of AstraZeneca's Iressa.

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

FDA AND ASCO WORKSHOP ON LUNG CANCER ENDPOINTS

Alexandria, Virginia
April 15, 2003

The FDA and the American Society of Clincial Oncologists (ASCO) jointly sponsored this workshop on lung cancer endpoints. On the 18-member panel were five FDA officials, five oncologists (including the president of ASCO), two statisticians, two patient advocates, two pharma officials, and two National Cancer Institute officials. There were surprisingly few people in the audience, and sources speculated this was because people may not have realized the session was open to the public.

From ASCO's perspective, the purpose of this meeting was to create a white paper with some suggestions about endpoints for approval of drugs for lung cancer. ASCO intends to submit that document to ODAC for its consideration, with the hope that the panel will forward something to the FDA as an advisory document. In the future, similar white papers are planned in other cancers, including breast, colon and perhaps hematologic cancers.

From the FDA's perspective, this meeting was not a replacement for, or duplication of, an ODAC panel. It also was not an advice-giving meeting because the FDA can only take advice from ODAC. Rather, the meeting was viewed as a forum to discuss the pros and cons of various lung cancer trial endpoints. Dr. Richard Pazdur, Director of Oncology Drug Products for the FDA, said, "This is hopefully one of many meetings. This is somewhat of a trial. We have not done this in the past. If it goes well, we will duplicate it in other diseases…And before we take anything to ODAC, we will have several of these meetings." Another FDA official said there are likely to be two more similar meetings this year.

This was not a meeting about AstraZeneca's Iressa (gefinitib, ZD-1839) or any other specific agent, Dr. Pazdur emphasized. He said, "This also is not about any specific drug under consideration at the FDA. We would not apply retrospective discussions and try to make some decisions about ongoing applications. I want to be clear that applications under discussion at the FDA or under review, commitments we've made and prior discussions we've had with sponsors will continue, and we will honor those commitments. This is not about a specific drug." Underpowered trials have become a major problem for the FDA. Dr. Pazdur said, "We are gravely concerned about underpowered trials…The number of patients is difficult to estimate, and many times the number is chosen on the practical basis of how many patients you can treat. This is further compounded when only one trial is coming in. Sponsors are not estimating sample sizes on true beliefs but on the practicality of how many patients they can get in the shortest time…and one trial becomes very problematic…One of the reasons we are having only one trial come in is that this is a crap shoot for many people (sponsors). When we are developing hormones, we have little problem asking the company to do two trials; companies are more

than happy to do that. They have a level of security that hormonal therapy will work in breast cancer. But when they are working on cytotoxics, there is a gamble companies face."

During the last 13 years (January 1990 to November 2002):

- 55 oncology drugs were approved by the FDA
  - 18 based on a survival endpoint
  - 1 based on time-to-progression (TTP)
  - 26 based on response rate
- 73% of all FDA oncology drug approvals were ***not*** based on a survival endpoint
- 67% of approvals were ***not*** based on survival when accelerated drug approvals are excluded

The three key issues at this meeting were:

- Can TTP be used as an endpoint? **No, there is no consensus whatsoever on this issue yet.**
- What is a non-inferiority trial? **Not directly answered, but approvals probably will require stronger data, and a p-value much better than p<.05.**
- Are patient reported outcomes (PROs) ready for prime time (translation: primary endpoint)? **No, but there is interest in working towards use of PRO as an end-point, and it could be useful as a secondary endpoint.**

A Canadian study suggests that stable disease (SD) correlated with partial response (PR), and an ECOG analysis found the same relationship. Data on this is expected at ASCO 2003.

THE INTERNATIONAL PERSPECTIVE

In the U.S., a randomized trial with overall survival as the endpoint is the accepted standard for FDA approval. In Europe, regulators also usually require a randomized trial, but other endpoints (response rate, TTP and quality of life) and other methodologies are accepted. Japan is the most lenient in approvals, usually not requiring randomized trials. Furthermore, when regulatory agencies in different parts of the world look at the same data, they may come to different conclusions. Dr. Pazdur said, "Japan for many reasons is still an outlier in drug approval…They haven't had a good infrastructure to do randomized trials until recently…and they want Japan-only data, thinking there is an ethnic difference…They are more (concerned with) safety than confirmation of efficacy."

There is no coordinated international review of drugs. Dr. Pazdur said, "I have disappointment over the lack of communication with international regulatory bodies. We talk in generalities and blue sky approaches, but when it comes down to specific applications there is very little communication. We don't call EMEA or Japan and ask what they are doing on something. We are attempting at least to discuss endpoints. At ASCO (annual meeting in May 2003) we will have an international drug regulatory meeting in open session and then closed meetings. I think we should invite members of EMEA, Japan, and Canada to these meetings. It is important that we move in one direction. Drug development is going on globally, and for us to put on blinders or for companies to get conflicting information is quite disconcerting and counterproductive."

Comparison Of Trial Endpoints

Endpoint	Opinion	Advantages	Disadvantages	Comments
Survival	Gold Standard. Most commonly used to determine efficacy of a particular regimen.	Guarantees efficacy if superiority design and beats anything. Easily determined. Hard to "fudge."	Crossover – can cause loss of demonstrated survival benefit	Non-inferiority designs are problematic with current regimens
Tumor response rate	Objective response (OR) is well-defined and widely accepted	Can be assessed in single arm study. RR correlates with symptom benefit.	Only documents activity in a subset of patients or does not correlate well with overall survival	OR may be more useful if stable disease (SD) included
Time-to-progression (TTP)	Poorly defined. Could be useful when there is long time lag between death and progression	Can be measured in all patients. Can be used with cytostatic drugs. Is assessed before crossover. Requires smaller studies	Not standardized. Is indirect measure. Clinical meaning is unclear. Can be expensive to measure carefully.	Does it measure clinical benefit? Is it reliable? It is not clear that because TTP improves that survival improves, nor is it true that because TTP increases that symptoms lessen. There is a rough correlation with overall survival (median survival is roughly twice the TTP)
Tumor-related Symptoms	Evaluation of patient morbidity has supported FDA approvals.	---	Lack of blinding, missing data,	Frequently discussed but not yet successfully used
Percent progression at a defined time	Not a traditional endpoint	May correlate with overall survival	Association with patient benefit unknown	Use implies that SD patients have a survival outcome similar to those with tumor regression.

Small Cell Lung Cancer Drug Approvals Around the World

Drug	U.S.	Europe	Japan
Etoposide	Yes	Yes	Yes
Carboplatin	No	Yes	Yes
Doxorubicin	Yes	Yes	No
Etoposide phosphate	Yes	Yes	No
Topotecan	Yes	No	Yes

ACCELERATED APPROVALS

Dr. Pazdur emphasized that accelerated approvals cannot be based on borderline evidence. He said, "You still must have substantial evidence. For example, we have had studies with a trend in survival only and no other evidence of benefit. That is a trend, not substantial evidence…It may seem paradoxical that we would accept a single arm trial for response rate for accelerated approval, but that can be substantial evidence. We believe you can tell response rate with a single arm study."

The take-home messages from the ODAC meeting in March 2003 on accelerated approvals, according to Dr. Pazdur, are:

- Confirmatory studies should be part of the drug development plan.
- There should be early discussion of confirmatory studies with the FDA.
- ODAC wants to be consulted on confirmatory study plans.

The FDA has the option of accelerated withdrawal for drugs, which get accelerated approval and then fail to show clinical benefit in confirmatory studies, but FDA officials emphasized that this really is not a politically viable option. Dr. Pazdur said, "My personal feeling, not the FDA policy, is that it will be very difficult to remove a drug from the market once it (gets accelerated approval) because obviously the Phase IV is only one part of the drugs history…but we are continuing to meet publicly at ODAC and privately with those sponsors to review their plans. The major thing the FDA can do is be prospectively active in designing protocols. They should be part of a comprehensive drug development plan. For those already out there, there is little we can really do other than exposing them to the light of day. The possibility always exists they could be taken off the market, but that is difficult. Drugs come off for toxicity, but for lack of efficacy it would be somewhat difficult. I'm not saying we can't and that it won't happen, but it is difficult."

The FDA has several approaches it hopes will address this problem, including**:**

- The FDA is looking at bringing the National Cancer Institute into the post-Phase II meetings it has with sponsors.
- Negative post-approval data may be included in labeling and advertising.
- The agency plans to start requiring that a Phase IV plan be submitted with an accelerated approval application. Dr. Pazdur said, "We think that is as important as the accelerated approval basis. If someone sees positive Phase II data, it will take time to get the information ready to submit, and the FDA has six months to review it. In that time a Phase IV could be negotiated and potentially started. We want to see a good faith effort…Why not start studying (combinations) earlier in the course of approval rather than waiting until the approval and then studying it? Let's be a little more responsible in addressing these issues."

The ODAC panel also has its own answer: take a tougher stance on accelerated approvals. Thomas Fleming PhD, Chairman of the Department of Biostatistics at the University of Washington and a frequent FDA consultant to ODAC and other FDA advisory committees, said, "(The ODAC panel meeting) left some of us with the feeling that accelerated approval is basically full approval, and if there isn't going to be accelerated withdrawal, then there is higher bar for endpoints for accelerated approval…I'm more willing to be lenient in approval if there will be accelerated withdrawal…I don't think there has been an accelerated approval in lung cancer so far…so, with me, that bar goes higher because it is tantamount to full approval."

NON-INFERIORITY TRIALS

A workshop participant thought the FDA was taking a very conservative approach toward non-inferiority trials, and he wondered if that was likely to change, but other participants pointed out that non-inferiority trials are not easy to interpret. Dr. Fleming said, "The bottom line is that non-inferiority trials can be very problematic in interpreting the results." ASCO president Dr. Paul Bunn, Director of the University of Colorado Cancer Center, called non-inferiority trials "tricky" studies."

Interpreting non-inferiority trials can be "touchy" and is a major problem in lung cancer, Dr. Pazdur said, so the FDA may bring in statistical consultants to help ODAC. He explained, "In lung cancer, most sponsors are coming in with one randomized trial now…so we will have to base our judgments on minimal databases from historical data…I see that as a big problem for non-inferiority trials in the future. Sponsors are not willing to do two randomized trials…and that affects the totality of the evidence…what we are looking for in a second trial is duplication of results. If someone is coming with one trial, the statisticians want a greater level of confidence – p=.05 squared (p=.0025) – but we haven't gone thatfar…We could spend a whole conference on non-inferiority. Guidance needs to come out on this whole issue

from our statistical group, but we want to see a prospective plan rather than a retrospective analysis."

TIME TO PROGRESSION (TTP)

There are numerous problems with TTP (which on average is four months in lung cancer) as a primary endpoint, and the workshop members had a long and lively debate about using TTP, with no resolution of the issue. A statistician said, "There could be an advantage to TTP as a surrogate if it is validated because you get a signal sooner in smaller numbers." An oncologist said, "It seems to me better to wait (for survival data)…I think we are more likely to get wrong conclusions if we use this (TTP)." Another oncologist said, "We have more drugs approved in lung cancer than any other disease than breast cancer and with less impact on survival. There is no curative chemotherapy in lung cancer, but we do have it in breast cancer. If you want to focus on patient symptomology, focus on that. I don't think you need a way to measure disease progression in this disease other than what you already have."

Among the problems that were identified:

- **Bias.** An expert said, "Even if TTP is measured at a fixed point in time, if progression took place prior to that, that is a bias." An FDA official added, "Except you know the bias is being applied equally."
- **Lack of blinding.**
- **Measurement.** There was no agreement on when the measurements (e.g., 6 weeks, 4 months) should be taken and how often they should be taken. An NCI official said, "The NCI would prefer TTP at six weeks…There is a 50% rate of progression for inactive agents at six weeks." Dr. Bunn said, "There are fewer deaths confounding it at six weeks than at four months." A statistician wondered, "But is there enough signal at six weeks that you get events?...That may be too short…Since death unfortunately is realized early (in lung cancer), what is the advantage of approving a drug on TTP if death occurs not long after?" An FDA official said, "We feel a lasting benefit would be more likely to be seen at four months."
- CT scans are often obtained every other cycle.
- Variable cycle lengths.
- Crossovers.
- High degree of subjectivity.
- Lack of sufficient data on the validity of this endpoint.

Dr. Richard Pazdur summed up the debate, saying: "I'm hearing that TTP is not ready for prime time and needs to be studied more. It may have more utility in diseases with longer survival…At this time we are not ready for this endpoint." Dr. Grant Williams, Deputy Director for the FDA's Division of Oncology Drug Products, Office of Drug Evaluation and Research I, added, "The discomfort with TTP is with TTP measurement and the challenge is to agree on the best way to do that and evaluate it. There isn't a lot of comfort with TTP as an endpoint, but we should look for the best way to validate it and cross-validate it." In line with this, ASCO wants to come up with a plan for how to proceed to collect the data needed to make TTP a surrogate marker in lung cancer, but it will not be easy and it won't be soon. Dr. Bunn said, "I would like to see the white paper have a plan for how to proceed (on this). People are intrigued with the potential for progression as a surrogate, so we need a plan on how that might be done, even if we say it is not ready yet, but we could say it is worthy, and here is a plan for how to do it."

PATIENT REPORTED OUTCOMES (PROS)

Dr. Richard Gralla of the New York Lung Cancer Alliance argued in favor of the use of PROs, suggesting quality of life should become the most used parameter in clinical trials. Workshop members were more receptive to this than TTP, but they also did not think this is ready for prime time.

There are three instruments currently available to measure quality of life in lung cancer, all of which have been validated in up to 20 different languages:

- Lung Cancer Symptoms Scale (LCSS)
- EORTC
- FACT-L

The issues that come up with the use of quality of life measures in Phase II trials were described as:

- Lack of investigator commitment
- Cumbersome instruments
- Patient deterioration
- Palliative care

The issues that make PROs difficult from a regulatory perspective include:

1. **Missing data.** Dr. Fleming said, "The reality is that people with poorer quality of life may be systematically lost to follow-up because of death or deteriorating quality of life…The only way to deal with this is to prevent it." Members generally agreed that defining missing data would help – and that should include saying that patients who die are not missing data.

2. **Blinding.** In oncology a large percentage of trials are not blinded because of drug delivery or toxicity issues. Dr. Fleming said, "If the effect is sufficiently profound that it exceeds what is due to unblinding, then I would find it convincing, but to sponsors that means there is a higher bar…I think unavoidably there has to be some raising of the bar (with use of PROs)."

3. **Composite endpoints.** Dr. Fleming said, "The thing that bothers me the most is aggregation of multiple domains, a weighted-average score. I want to be confident that if there is a change in the average it reflects a change in domains. If we put relative values on the domains, that makes sense…It makes sense to have patients identify their important domains. But if you do that and the intervention is only effective against that, and that is something important to only 20% of patients, then this has less clinical relevance."

4. **Drug mechanism.** There is a concern that a drug designed to treat lung cancer could make patients feel better by relieving pain, for example, but affect the course of the disease.

5. **Exploratory analyses vs. confirmatory results.** Dr. Fleming explained, "Exploratory results can't be taken the same way as confirmatory results. The need for control is inherently important."

6. **Power and sensitivity/specificity.** A statistician explained, "I think power is a key issue…the issue of power is inherently related to whether the instrument will be sensitive to pick up a clinically meaningful effect." Another statistician said, "We just finished a large emphysema study with dyspnea as the endpoint…and our investigators said quality of life should be the primary endpoint, but it wasn't because of considerations of power. Instead, it was powered for survival. Quality of life was the least sensitive measurement." An oncologist said, "Quality of life measurements are noisier and because of that sometimes power calculations designed for other endpoints don't have the power (you need)."

   An FDA official wondered, "If one or two symptoms get better and others don't get worse, is that sufficient? Or, is it power problem?" A statistician answered, "It could be power problem."

7. **Mean results vs. individual results.** At the end of the debate, ASCO's Dr. Bunn asked the FDA officials if there are instruments valid enough that a drug could be approved on PROs. Dr. Pazdur replied, "It is theoretically possible, but when we give advice to companies we also are looking at a real world scenario…Is it an achievable situation?…Will you have 90% data, will you eliminate the bias? We are aware of the problems. Will a company be able to overcome them? Granted, if there were a positive primary (PRO) endpoint, and all the negative factors were minimized, you would interpret that in a positive light, but how often will that happen? That is the cold hard facts of life…Would a company really feel comfortable, given those hurdles in the existing world, in making this a primary endpoint for a trial?" A pharma official in the audience shook his head ***no***. Dr. Gralla concluded, "We are not ready for this if we insist on non-inferiority (as well). Then, I think PRO is important, can support an application, and should be looked at…and if you do have non-inferiority, you will get approval anyway." Dr. Bunn commented, "I'm not sure companies will come in with PROs at this point."

ENDPOINTS IN ADJUVANT AND NEOADJUVANT STUDIES

The ASCO president said there would be some data at ASCO 2003 on endpoints in adjuvant and neoadjuvant studies. He wondered whether TTP is a relevant endpoint for these trials, noting that some melanoma drugs have been approved on this basis. The FDA's Dr. Williams said, "Disease-free survival has been acceptable in the past for tumors that are symptomatic – primarily breast cancer. We don't have a firm policy." An NCI official said, "It is hard to make a determination that disease-free survival correlates with overall survival because there are no trials that show that. In my own opinion, I think it is fine to assume that, but it relates to the toxicity of the therapy." Dr. Pazdur indicated that TTP might be an acceptable endpoint in these trials as they wait for survival data in a Phase IV trial."

MISCELLANEOUS ISSUES

Trial design: Many accelerated approval requests in oncology have been in the refractory setting, with single-arm trials. Dr. Pazdur explained some of the issues this raises for the FDA: "Toxicity is difficult to analyze, and we can't look at time-to-event endpoints. We also are facing a problem of how low a response rate could be. What is the minimal cut off that one could expect? My biggest concern is that, as we develop a more refractory disease population, are we looking at artificial and niche populations where the response rate may not apply to a more general population? We try to encourage sponsors to do randomized trials, looking at an interim analysis and then going forward for a demonstration of clinical benefit through survival. I think that is a cleaner way, but we will meet some resistance because it is more expensive…It won't do away with single arm trials – they are still a viable option – but they are problematic…After one panel *(NOTE: He probably is referring here to Iressa)*, at the next couple of sponsor meetings, people asked, 'Can we get our drug approved on an 8% response rate, on 5%, on a database of 34 patients?'"

Combination trials: A workshop member said, "What I see coming are targeted agents that don't block just one pathway…or that need to be given in combination…To what extent does the sponsor have to prove that both drug A and drug B are needed?" Dr. Pazdur said that both drugs have to show they contribute a benefit, "The strongest evidence is a clinical trial with three arms. If you want to depend on pre-clinical data, we need great deal of assurance that it was real, reproducible, etc. And you know that preclinical efficacy in

oncology has been 'rather tenuous'…we are working with rather marginal drugs here."

Gene mutations: Asked if the FDA is ready to approve a drug based on efficacy in patients with a particular gene mutation, Dr. Pazdur said, "Hypothetically. If there is sound scientific rationale to support that, yes. We have defined subsets before. There is no difference between labeling for a genetic subset vs. an histologic subset…And there has to be buy-in from the scientific community, but there is not a regulatory barrier."

CMS reimbursement: ASCO president Dr. Paul Bunn expressed concern that CMS may stop paying for some chemotherapy drugs and wondered if CMS had the statutory authority to deny coverage for drugs considered safe and efficacious. Dr. Pazdur pointed out that there are European countries where approved drugs are not reimbursed. Another workshop participant said, "It appears CMS is backing off (on oncology drug coverage, including off-label coverage). ASCO and patient groups have written letters that this is illegal…if off-label uses have to be covered, then FDA-approved uses should be covered."

***AstraZeneca's Iressa*:** There was some oblique but open discussion of Iressa, which is unusual for a drug currently under FDA-review. A NCI official said, "I have a question about Iressa on quality of life. Were you suggesting the tool was inappropriately applied or the tool was inappropriate? Are there Phase II situations where quality of life can give meaningful measures in terms of drug approval or are we really talking about Phase III comparative trials?" Another workshop participant said, "The tool has some problems, but if it didn't then in Phase II, there is so much confounding data coming in...It looks interesting but without a control you don't know how interesting. And (in third-line disease) there isn't even a historical comparison…(But) it is interesting to see in any response (in that population)."

A workshop member, asked about Iressa after the meeting, said he believes it is effective. He was involved in the clinical trials, and he believes the findings in the Phase II IDEAL (1 and 2) and the Phase III INTACT are all true but unrelated. That is, the negative findings in INTACT should not tarnish the positive findings in IDEAL. However, he expects Iressa to be a "niche" product that only helps a small percentage of patients. He also indicated ***there will be clinical phenotype data at ASCO 2003 that helps identify which patients will respond to Iressa***.

♦