Progression-Free Survival Remains Debatable Endpoint in Cancer Trials

By Rabiya Tuma

In recent years the U.S. Food and Drug Administration has approved cancer therapies on the basis of improved progression-free survival (PFS), including several drugs for the treatment of metastatic kidney cancer. Clinicians who treat kidney cancer patients are convinced that the agency made the right move, and some even contend that overall survival (OS) is no longer a useful endpoint in this disease setting.

However, despite some support for the use of PFS in one disease setting, its value as a standard endpoint for clinical trials remains a point of debate in oncology as a whole. Some experts argue that OS is the only time-to-event endpoint that shows true benefit to patients, whereas others think that PFS should be accepted as long as the trials are carried out according to rigorous standards. To hash out some of those issues, academic and industry statisticians gathered Oct. 7–9 for the Progression-Free Survival Oncology Workshop in Bethesda, cosponsored by the Drug Information Association, the FDA, the National Cancer Institute, and the Pharmaceutical Research and Manufacturers of America (PhRMA).

Richard Pazdur, M.D., director of the FDA Office of Oncology Drug Products and the keynote speaker at the meeting, acknowledges that PFS requires a more nuanced risk–benefit analysis than OS does, but he says that sometimes PFS is an appropriate endpoint for registration trials. “A lot of it has to do with magnitude, magnitude, magnitude,” he said in an interview. “Just demonstrating a statistically significant difference in PFS is not enough. It has to be clinically meaningful.”

And he made it clear that even as the agency considers more approvals that are based on PFS, it won’t abandon OS altogether. “We are always still interested in looking at OS and will look at it because OS is not just an efficacy endpoint but also a safety endpoint.”

The Case for PFS

In less than 4 years, the FDA has approved five agents for the treatment of kidney cancer. Only one showed a statistically significant increase in OS compared with interferon, which was the previous standard of care. Three of the other drugs were approved based on PFS and the fourth on overall response rate (it later showed a benefit in PFS). All showed a nonsignificant trend for improved OS in the registration trials, though. When asked about this pattern, Pazdur said that the agency and the trial sponsors were concerned, on the basis of strong phase II data, that a large proportion of patients initially assigned to the interferon arm would cross over to the experimental agent when their disease progressed, and that this crossover would confound any OS analysis.

“Hence, we said if there was a robust effect on [PFS], then we would consider it a registration endpoint designating clinical benefit in this patient population,” Pazdur said. “This situation reflects a basic problem we have in measuring [OS], in that it can be confounded by subsequent therapies and most important by crossover to the experimental drug, either in the trial or off trial, if it is available on the market.”

In the most recent example, clinical trial investigators conducted the final OS analysis for two phase III clinical trials comparing bevacizumab plus interferon with interferon alone. At the annual meeting of the American Society of Clinical Oncology this year, they...
reported that the large, statistically significant gains in PFS did not carry through in the form of a statistically significant increase in OS. (Bevacizumab plus interferon gained FDA approval in July 2009.)

Nicholas J. Vogelzang, M.D., chair and medical director of developmental therapeutics at US Oncology Research, who discussed the papers at ASCO, said that he was mildly disappointed that the survival analysis was not positive, but he stressed that there was a trend toward a benefit in OS. Perhaps, he suggested, OS is no longer a valid endpoint in kidney cancer.

“We do want increases in OS,” Vogelzang said in a recent interview. However, with so many active agents available to patients after they progress on the study drug (or cross over to the experimental drug arm), increases in survival will be hard to detect in trials. “PFS is a strong enough endpoint,” he continued. “It is the only reliable consistent endpoint that we have right now. You can hem and haw, but ultimately PFS has now become the answer to the question, ‘Is this drug better than another drug?’” Vogelzang has received honoraria from and consults for several companies, including Genentech, maker of bevacizumab, but does not own stock in any of the companies.

The lack of OS improvement makes sense, according to Marc Buyse, Ph.D., executive director and biostatistician at the International Drug Development Institute in Louvain-la-Neuve, Belgium. If the trials could be conducted without allowing patient crossover, then the survival benefit should be detectable, but that has not been the case. “The crossover design is used for ethical reasons, though it creates complete confounding in terms of survival,” he said. “If we have very active new drugs—and I think that it is clear that the kinase inhibitors are very active for metastatic renal cell cancer—and if they are tested with these kinds of crossover designs, then it is hardly surprising that there is no benefit in survival. The paradox is that the more active and the better the drug is, the less survival will be affected in such a design.”

Brian Rini, M.D., a staff member in the department of solid tumor oncology at the Cleveland Clinic Taussig Cancer Center in Ohio, who led one of the bevacizumab-plus-interferon trials, is confident that the new agents are effective and providing patient benefit, even without statistically significant OS endpoints in the clinical trials. “The median survival is now about 2.5 years,” said Rini, who has received research funding from Genentech and consults for the company. “With interferon, it was about 1 year. That didn’t happen by magic; it happened because there is active drug. To me there is no question about it. Can we say with confidence that drug X makes you live this amount longer and drug Y that amount? No, we can’t, but I don’t know that that is that critical. If patients are well enough, they are going to receive multiple sequential therapies, and it is really the cumulative effect of those therapies that is extending their lives.”

The Debate

But not everyone is ready to give up on OS. Tom Fleming, Ph.D., professor of biostatistics and statistics at the University of Washington, Seattle, concedes that in rare cases, including some of the kidney cancer trials, PFS is acceptable. “There is a line that you can cross where the effect [on PFS] is so enormous that you can judge that a survival benefit has effectively been shown,” he said.

But Fleming is adamant that crossovers shouldn’t be allowed in a trial. “Cross-ins to agent X shouldn’t happen. This is an agent that is experimental. It is not available unless you provide it to patients, and you shouldn’t provide it to them in a clinical study that by its nature is trying to find out if patients should even take agent X,” he said. Instead, patients should receive drugs that have already been shown to be safe and effective. In that situation, “if any survival benefit is diluted away, then it means that the experimental agent was not needed to improve survival in these patients. I am not saying that is what is going to happen, but if that is the excuse for not using survival, then the fact is that survival is still the right answer. It just means that your agent isn’t providing any net affect on survival in the context of what patients could achieve with other therapies,” he said.

William D. Bushnell, group director of oncology biometrics and epidemiology at GlaxoSmithKline and cochair of the PhRMA working group on PFS, agrees that theoretically that assertion is correct. “It is true that if you had a huge difference in PFS, it should translate into an OS advantage. But given the setting where we are today, with numerous examples where we have a PFS difference and a much smaller effect on OS, we know it is hard to show an OS advantage.” If patients in a trial have a median PFS increase of several months it may not be realistic to expect that difference to carry over to OS, which can last several years beyond progression.

Bushnell, who is also a member of the program committee for the PFS workshop, thinks it would be good for trial sponsors to monitor patients more closely after progression and track what other therapies they use between progression and death. But he points out that a key advantage of PFS is that it is a purer comparison of the effect of the therapy being studied.

Fleming and the NCI’s Boris Freidlin, Ph.D., another member of the program committee for the PFS workshop, contend though that showing an increase in PFS alone does not constitute clinical benefit. “If the tumor grows more slowly but the patient dies at the same time, and there is no reduction in pain or something like that, it is not clear what the benefit is here,” Freidlin said. “Remember, PFS is progression defined totally arbitrarily: 20% increase in measurement of longest tumor dimension, not some significant clinical deterioration. And it is a relative measure of 20% for both large and small tumors.”
Pazdur is well aware of the debate surrounding PFS and clinical benefit and used his opening remarks at the workshop to raise the question of how much PFS needs to increase to provide a clinically meaningful benefit for patients. In an interview before the meeting, he explained: “If we are using PFS because we believe there is a confounding of survival, we need to ask what would one expect to see in OS and how much bigger the effect would have to be in PFS to get that. That increase in PFS is what we should be aiming for,” he said. “We always see [trial] sponsors coming in with a lesser and lesser effect or proposed effect on PFS that they would like to demonstrate. We have stated that even a small effect on OS is of clinical benefit that does not translate to an analysis of PFS.”

To get an idea of how large an increase in PFS would have to be gained in a trial to translate into improved survival, Buyse’s team performed a meta-analysis on patient-level data from phase III colorectal and breast cancer trials. They found that a 20% increase in PFS in colorectal cancer predicts a statistically significant increase in OS. However, they could not identify a threshold for PFS improvement in breast cancer that predicted a statistically significant survival benefit. That outcome implies, he said, that PFS is not a good surrogate for survival in that disease. In fact, PFS may be a valuable endpoint in some disease settings and with some drugs but not in other cases. His group aims to perform a similar analysis in kidney cancer, and members are collecting the patient-level data for that project.

Building Better Consensus

Both Bushnell and Pazdur said they hope that the workshop will lead to more uniformity in how missing data points, such as missed scans or x-rays, are handled, as well as to a better agreement on when and how independent review committees are used. In general, the FDA requires such committees when trials are unblinded, but Pazdur also thinks that they are necessary when toxic effects in one treatment arm or another could lead to partial unblinding.

As if demonstrating that consensus on these topics won’t be easy to reach, Bushnell said he thinks that partial unblinding is not as concerning as some regulators and researchers make it out to be. “That is a worry if 100% of patients in one arm have an adverse event and none on the control arm do. That is never the case,” he said.

Bushnell said the PhRMA working group is presenting data on these issues, as well as on new methods developed to allow for auditing a subset of patient data, instead of using an independent review committee to reevaluate all data in the trial.

“This is a very difficult area and one full of emotion,” Pazdur said. “We have been working on this area for the past almost decade since I came to the agency . . . it is kind of a circular argument you get into. There are people who say you absolutely need OS, and there are people claiming that PFS has a meaning and you are being overregulatory by demanding OS.”

As for the recent approvals of kidney cancer therapies that are based on PFS, he said that he has no second thoughts.