Trends-in-Medicine

January 2009
by Lynne Peterson

Quick Pulse

Trends-in-Medicine
Stephen Snyder, Publisher
2731 N.E. Pinecrest Lakes Blvd.
Jensen Beach, FL 34957
772-334-7409 Fax 772-334-0856
www.trends-in-medicine.com
TrendsInMedicine@aol.com

FDA Mulls Genomic Biomarkers

Gaithersburg, MD
December 17, 2008

The FDA's Oncologic Drugs Advisory Committee (ODAC) met in December 2008 to discuss the use of genomic biomarkers in cancer trials. The panel was briefed specifically on the use of K-ras for Lilly/ImClone/Bristol-Myers Squibb's Erbitux (cetuximab) and Amgen's Vectibix (panitumumab) in colorectal cancer (CRC), but the purpose of the meeting was to look more broadly at the regulatory path for genomic markers in general, not just K-ras.

The panel appeared convinced of the value of K-ras for predicting response to Erbitux and Vectibix in CRC, though that was not the issue before them. No votes were taken, but the panel provided the FDA with lengthy discussions of several issues of concern to the FDA. Members told the FDA that:

- Limiting a drug to a subgroup of patients based on a retrospective biomarker analysis depends on each individual situation. It may make sense where the assay lets a patient avoid toxicity but could face patient/physician resistance if it removes patient hope.
- A prospective study may not be necessary if there are very good, multiple retrospective/prospective trials with sufficient tissue samples. It is unrealistic to expect that biomarkers can always, or even often, be identified before a drug trial begins.
- A consistent effect on progression-free survival (PFS) is probably sufficient, even in the absence of a demonstrated effect on overall survival (OS), though the importance of OS cannot be minimized.
- Guidelines don't necessarily apply to biomarker development, provided data dredging isn't used to identify the biomarker, too many variables aren't involved, and a single biomarker is being studied.
- Limiting future trial enrollment based on a biomarker is tricky and should not be done prematurely.
- Biomarker tests need to be used the way they were studied in trials.

Recently, the American Society of Clinical Oncology (ASCO) recommended that K-ras testing be required before the use of an EGFR inhibitor, and researchers estimated that K-ras testing of colorectal cancer patients could save hundreds of millions of dollars (*See page 10*).

The FDA Perspective

At the beginning of the panel meeting, Dr. Richard Pazdur, director of the FDA's Office of Oncology Drug Products, Office of New Drugs, Center for Drug

Evaluation and Research (CDER), outlined the FDA's dilemma with biomarker testing. He said, "The selection of a drug based on genomic biomarker profile is desirable because it limits drug exposure to patients who will benefit or are most likely to benefit from drug treatment, avoids drug use in patients who will be or are likely to be harmed by drug treatment, or enhances safe use by optimizing drug dosing. In the ***ideal*** case, the development of the assay methodology for the genomic biomarker should be an integral part of the clinical drug development program, such that the clinical studies required to establish the efficacy of the drug and those needed to establish the prognostic and predictive value of the genomic biomarker as measured by a well-characterized assay occur ***in tandem***. FDA strongly endorses such scientifically-guided drug development as part of FDA's Critical Path Initiative… This ideal approach to drug development continues to be underutilized."

Dr. Pazdur said the FDA is concerned about "retrospective" or post-hoc genomic biomarker assessment of clinical disease characteristics, "In the worst examples, this involves a retrospective *re-analysis* of a 'failed' clinical trial in which efficacy is purported to be established in a subset defined by a genomic biomarker/patient characteristics without consideration of multiplicity (i.e., data dredging), substantial missing data, and poorly characterized assays. FDA discourages such practices." He added that these types of studies should not be part of the panel's discussion.

Yet, the FDA recognizes that there may be legitimate reasons for failure to prospectively consider the impact of genomic biomarkers early in drug development due to advances in the scientific knowledge of a drug or disease that occur while drug development is ongoing. Dr. Pazdur todl the panel the FDA was seeking its guidance regarding how to incorporate new scientific information without compromising the legal mandate to ensure that marketed drugs show substantial evidence of efficacy and are reasonably safe.

The levels of evidence needed may differ depending on the claim being sought. For example, restriction of drug use to patient subsets to improve safe use of the drug might not require the same level of scientific rigor as claims for specific drug benefits.

As a starting point and to put the discussion in context, the panel was given an example of real-world retrospective biomarker analyses intended to support changes to product labeling and support device approval – K-ras markers for Erbitux and Vectibix. Both ImClone and Amgen have proposed inclusion of information on drug use in the subset of patients with metastatic colorectal cancer (mCRC) whose tumors express wild-type (WT) K-ras. While there were specific discussions of K-ras, Dr. Pazdur emphasized that the real issue before the panel was more general, not limited to K-ras, Erbitux, or Vectibix.

Dr. Ruthann Giusti, a medical officer in the FDA's Division of Biologic Oncology Products, Office of Oncology Drug Products, Office of New Drugs, CDER, reviewed the regulatory history of genomic biomarkers and K-ras. She emphasized that the FDA believes that the optimal approach would be to conduct an adequate and well-controlled trial, ***prospectively*** designed to assess efficacy in subgroups based on K-ras testing by a validated assay. However, the widespread publication and presentation of the retrospective K-ras analyses have resulted in practice changes in the community. Thus, ***prospectively designed trials may no longer be feasible.***

Given these practical considerations, the FDA indicated that retrospective analyses from clinical trials could be submitted provided that:

- The trial was adequate, well-conducted, and well-controlled.
- The sample size was sufficiently large to be likely to ensure random allocation to each of the study arms for factors (i.e., K-ras status) that were not used as stratification variables for randomization.
- Tumor tissue was obtained in ≥95% of the registered and randomized study subjects and an evaluable result (wild-type or mutant K-ras) is available for ≥90% of the registered and randomized study subjects.
- Prior to analysis, FDA has reviewed the assay methodology and determined that it has acceptable analytical performance characteristics (e.g., sensitivity, specificity, accuracy, precision) under the proposed conditions for clinical use.

K-ras Randomized Trials

Trial	Line	Additional therapy	Number of patients	% of ITT	Assay	Trial primary endpoint	Hazard ratio K-ras WT vs. Mutant	
Erbitux								
CRYSTAL	1st	FOLFIRI	540	45%	PCR-based	PFS met, p=0.048	0.68	1.07
NCIC CO.17	3rd	BSC	394	69%	Sequencing	OS met, p=0.0046	0.55	0.98
EPIC	2nd	Irinotecan	300	23%	Sequencing	OS not met, p=0.71	1.29	1.28
OPUS	1st	FOLFOX	233	69%	PCR-based	RR not met, p=0.06	---	---
Vectibix								
20020408	3rd	BSC	427	92%	PCR-based	PFS met, p<0.0001	0.45	0.99
PACCE	1st	Chemotherapy /Avastin	863	82%	PCR-based	Not met; ***inferior*** PFS, p=0.002	1.36	1.25

- Genetic analysis is performed according to the qualified assay method by individuals who are masked to treatment assignment and clinical outcome results.
- Prior to analysis of clinical outcomes based on the genetic testing, agreement with FDA has been reached on the analytic plan for hypothesis testing for proposed labeling and promotional claims.

Both Amgen and ImClone have proposed retrospective testing of K-ras status from large randomized trials that have completed accrual (Amgen) or are actively accruing patients (ImClone). The ongoing studies have been modified to enroll only patients with K-ras wild-type tumors through the completion of the studies.

Robert O'Neill PhD, director of the FDA's Office of Biostatistics, Office of Translational Sciences, CDER, discussed statistical issues in the design, analysis, and interpretation of biomarker use in clinical trials to support marketing or promotional claims. He said this is not just a K-ras issue but something the FDA is seeing across the board in a variety of diseases. He outlined the components of a good analysis plan for a retrospective evaluation:

- **Randomization** – Look for an overall treatment effect on the primary outcome in the ITT population (***everyone*** randomized). If there is a statistically persuasive result (usually $p<0.025$), then examine subgroups by looking at the marker negative group and the marker positive group and look for a differential treatment effect (an interaction test). If there is no statistical significance on the primary endpoint, then everything else is exploratory.
- **Marker status classification.**
- **Marker classification performance.**
- **Statistical control of false positive conclusions** – How many hypotheses, which were primary, and which failed? Also, the number of outcomes – OS, PFS, response rate (RR) – needs to be considered.
- **Data to generate the hypothesis vs. data to confirm the hypothesis.**

He cited several concerns with retrospective studies:

- The **classification factor** is not known at the time of the study initiation, so the study is, at first, not analyzed with that factor as part of the hypothesis.
- The **initial hypothesis and endpoints** for the study are not changed, except if pre-specified as part of a planned adaptive study design.
- The control of the **false positive** conclusions from the study are appropriately dealt with.
- The **randomization** is not stratified on a factor that itself is of interest as one of the hypotheses to be tested.
- The **factor of interest** is ascertained at baseline on all subjects randomized to treatment groups, and if not, what happens.

Dr. O'Neill offered a working definition of a prospective/retrospective study: "In completed or post-interim-analysis trial where genomic samples were collected prior to treatment initiation, whether or not full ascertainment, the genomic hypothesis is 'prospectively specified' prior to diagnostic assay testing. However, the clinical outcome data without genomic information have already been (partially) collected, unblinded, and analyzed. The genomic data analysis might be arguably 'prospectively' performed, which is a retrospective analysis."

Dr. O'Neill illustrated how there can be inconsistency between endpoints in the convenience sample, using K-ras in the EPIC trial as an example.

Inconsistencies in the EPIC Trial of Erbitux in CRC

Endpoint	WT K-ras	Mutant K-ras
Overall survival	HR 1.29	HR 1.28
PFS	HR 0.77	HR 1.00
Best response	10% vs. 7% (Nss, p=0.61)	12% vs. 5% (Nss, p=0.29)

He charged that the evidence of K-ras predictability is "not as consistent as you might think" from the Erbitux and Vectibix trials. He said, "I think there will be an important distinction between OS and PFS (of Vectibix by K-ras)…Neither of the P-mab (Vectibix) studies show any benefit to OS on K-ras." However, he admitted that the FDA does not have access to the K-ras level data, so their discussion is for "educational" purposes only.

LILLY/IMCLONE PERSPECTIVE

Dr. Hagop Youssoufian, senior vice president of clinical research and development at ImClone, a wholly-owned subsidiary of Lilly, noted that the dissemination of the K-ras data has been extremely rapid. In June 2008, the National Institutes of Health and the National Cancer Institute issued an action letter asking Erbitux investigators to suspend accrual for all CTEP-sponsored and Cooperative Group trials in CRC that included Erbitux in the treatment of CRC until appropriate modifications to the protocol and informed consent were made, new information concerning K-ras was added, and patients with mutated K-ras in tumors were excluded. The NIH/NCI letter came after the results of K-ras analysis from the CRYSTAL trial were presented at the American Society of Clinical Oncology (ASCO) 2008 meeting.

CRYSTAL met its primary endpoint (PFS), showing a significant improvement when Erbitux was added to FOLFIRI (in an intent-to-treat analysis) – 8.9 months vs. 8.0 months. The results presented at ASCO were based on a retrospective

analysis of K-ras status in subjects with tumor tissues evaluable for analysis of K-ras mutation and showed that the benefit of adding Erbitux to FOLFIRI was only demonstrated in subjects with K-ras WT in their tumors (hazard ratio=0.68, p=0.017). Subjects with K-ras mutations in their tumors derived no benefit from the addition of cetuximab over and above chemotherapy alone (HR=1.07, Nss, p=0.75).

Dr. Youssoufian reviewed the K-ras data on Erbitux from four randomized studies – NCIC CO.17, CRYSTAL, EPIC, and OPUS – in patients with mCRC and strongly supported a label change. ImClone contends that the data make a "compelling argument" that patients with K-ras WT tumors derive enhanced benefit from the addition of Erbitux to their therapeutic regimen while patients with tumors harboring K-ras mutations probably will not benefit from Erbitux or other anti-EGFR antibodies.

Dr. Youssoufian cited several reasons that support the use of K-ras as a predictive biomarker in CRC:

- Multiple independent studies with similar conclusions.
- Rigorous design of company-sponsored, well-conducted randomized studies
  - Positive results.
  - Prospective focus on a specific biomarker.
  - Prospectively-defined analysis plans for K-ras.
  - Blinded K-ras assessment.
  - K-ras evaluable subset representative of the ITT population.
- Consistent results independent of test methodology.
- Consistent results across Erbitux and Vectibix.
- NCI and cooperative groups use K-ras guided patient selection.
- Prospective trials to confirm the predictive value of K-ras are no longer possible.

While admitting that there are limitations to retrospective analyses, Dr. Youssoufian argued that the consistency of the results across the Erbitux studies strongly suggests that K-ras status is a predictive biomarker for Erbitux activity in mCRC. He said, "We believe the timing is right to discuss a label change" and to discuss inclusion of K-ras testing in the Erbitux label.

Ongoing Erbitux CRC Trials with K-ras Testing

Erbitux trial	Line	Design	K-ras testing prior to randomization	Randomized to K-ras WT only	Primary analysis K-ras WT only
N0147	Adjuvant	3,768 *	Yes	Yes	Yes
PETACC-8	Adjuvant	2,549 *	Yes	Yes	Yes
C80405	1st line	3,610	Yes	Yes	Yes
COIN	1st line	2,421	N/A	N/A	Yes
S0600	2nd line	1,260 **	Yes	Yes	Yes

* Initial trial size was smaller, but it was increased to this size.
** This is initial trial size, but an expansion is planned.

AMGEN PERSPECTIVE

Dr. Paul Eisenberg, global regulatory affairs and safety for Amgen, said the predictive value of K-ras has become clear and suggested that the use of K-ras as a predictive biomarker will improve the risk:benefit profile for Vectibix monotherapy in CRC patients.

- There is biologic plausibility of the K-ras hypothesis, and the activating mutations are well-characterized.
- The strengths of the monotherapy analysis include:
  - Demonstration of a high ascertainment of K-ras on archived samples (92%).
  - Pre-specified analysis plan was used.
  - A reliable assay was used.
  - A consistent high predictive value of mutated K-ras for non-response.

Dr. David Reese, global clinical development for Amgen, said the K-ras hypothesis in mCRC emerged in parallel with the clinical trial data, "Prior to conducting an analysis of our Phase III study we had to identify a K-ras assay…We sent samples to a variety of laboratories to identify an assay that could be used in routinely available clinical specimens." The choice was the DxS Mutation Test Kit. He sited several strengths of the Phase III trial (Study 20020408) in assessing K-ras:

- The protocol required tumor samples which were archived for potential biomarker correlative analyses.
- The expected K-ras evaluable sample size was sufficient to provide balance between the treatment arms.
- K-ras was the only biomarker evaluated, not a panel of biomarkers.
- There was >90% power to test whether K-ras was predictive for PFS.
- The statistical analysis plan was finalized prior to unblinding of K-ras status.

Ongoing Vectibix CRC Trials with K-ras Testing

Vectibix trial	Line	Design	Number of patients	Primary endpoint
20050181	2nd line	FOLFIRI ± Vectibix	1,187	PFS and OS
20050203	1st line	FOLFOX ± Vectibix	1,183	PFS

Utility of K-ras in Vectibix Monotherapy Trials

Vectibix trial	K-ras ascertainment	Objective response	
		WT	Mutant
20020408	90%	17%	0
20030194	96%	22%	0
20030167	91%	6%	0
20030250	84%	9%	0
Total	90%	14%	0

PANEL QUESTIONS FOR FDA AND INDUSTRY SPEAKERS

Consent issues: ***Are they solvable?*** **Yes.**

- *ImClone:* "In colorectal cancer…it is more feasible to get samples. In metastatic settings, tissue availability may not be the same as in the adjuvant setting…but in all Erbitux and panitumumab trials, EGFR testing was an entry criterion, so there was some tissue available, and that was the basis for going back and testing."
- *Amgen:* "Making specimen collection mandatory is one answer. Our goal is to have a sample from every patient in every trial. Our ascertainment rates were high – 92% in the pivotal trial…We have consent forms that apply across our programs."

Validation: ***How does the FDA validate or pass on a biomarker?*** **It has to show clinical relevance.**

- *Dr. Robert Becker, chief medical officer in the FDA's Office of In-Vitro Diagnostics, Center for Devices and Radiological Health (CDRH):* "It has to show safety and efficacy in some clinical context…The idea of clinical validation for the biomarker is central to a determination that the biomarker is appropriate for approval."
- *Dr. Patricia Keegan, director of the FDA's Division of Biologic Oncology Products, Office of Oncology Drug Products, Office of New Drugs, CDER:* "Since there is no approved biomarker...the arrangement we've reached is the analytical quantification had to be acceptable to CDRH, and once they reach that point, the analysis would serve to look at drug effect and at the predictive prognostic effect – so it would proceed together."

Future trial design: ***How will clinical trial design be impacted?*** **Larger trials are likely, not smaller ones.**

- *FDA's Dr. O'Neill:* "Some folks will claim that if you get it right and you actually do this assay testing early enough, that you get much more efficient trials later on… because you are enriching the population for a treatment effect…So, the downstream of early good work that characterizes the sensitivity and specificity of the classification strategy should pay off in terms of a smaller trial. The issue/concern is we've generally followed a practice of enriching trials…For example, in the cystic fibrosis area, you might take more sick patients because they are likely to have more events and require a smaller sample size…And you might extrapolate from that to a less sick population…Here the issue is to protect that group from toxicity, but you can get a more efficient trial…because the treatment effect will be greater. This goes to the impact of misclassification of marker status…Generally, equal class should drive the interaction effect more to a null or zero effect…So, it would be harder to detect a difference."
- *FDA's Dr. Pazdur:* "I think inherent in this whole process is a thoughtfully planned out process and doing your homework before entering into a Phase III trial. Unfortunately, that isn't the case. Many times we find a rush to a Phase III, and many times sponsors even wanting to skip a Phase II because they saw a few responses in a Phase I trial…This whole thing is based on a thoughtful approach…an understanding of the mechanism of action, a partner for *in vitro* diagnosis…That is one of my worries – an emphasis that yes we can go backwards, but there are problems with going backwards."
- *Panel member Richard Simon, D.Sc., chief of the Biometric Research Branch at the National Cancer Institute:* "In a situation like Herceptin (Genentech, trastuzumab for breast cancer), it leads to a smaller trial because people who benefit get into the trial…But in real life, developing a drug with predictive biomarkers will actually lead to larger clinical trials. It won't make life quicker, simpler, or cheaper, but it will be larger, more expensive trials because often you won't have full confidence in the diagnostic by the time you get to the pivotal trial…And you will need enough patients to test. And even in cases like Herceptin, where you think you really know who is likely to benefit from this drug, CDRH wants to have negative patients included anyway, and that raises serious issues for patients…You have to say, 'Here is a drug that we don't think helps you, but to show the FDA it doesn't help, the FDA wants you in the trial.'"
- *David Harrington PhD, a biostatistician from Dana-Farber Cancer Institute:* "This is a situation where science races ahead of trials…I agree with Dr. Simon that trials in practice will become larger…We will be aware of likely heterogeneous populations but may not be able to identify them when trials get started."

Tissue samples: ***Will tissue collection be a problem and what happens with patients without sufficient tissue?*** **Not all patients will have sufficient tissue, but off-label drug use is a possibility in those cases.**

- *Amgen:* "We require 20% of a section, but we can go to smaller amounts…You can get to <4 mm^2, down to potentially 1 mm^2…We haven't been able to determine on a per-cell basis the K-ras status…There are no tools to do that…The sensitivity is down to 5-10 copies of mutant K-ras." Asked what happens if enough sample tissue isn't available – is the drug denied or more tissue obtained – the Amgen official said, "Ideally, we would get additional tissue or, in the absence of that, the current label (for Vectibix) is for all-comers, and physicians can prescribe according to the current U.S. label."
- *FDA's Dr. Becker:* "CDRH is not interested in seeing trials accrue patients where there is subtle knowledge that patients cannot benefit from a drug, but to the extent that has not been settled definitively, then the opportunity to get to a definitive position is something we are interested in hearing ideas about...The idea at the end of the game is to be positive when a patient obtains a negative test and

says, 'Gee, doc, why can't I take this drug?'…Not that, 'It is our best guess that you won't benefit.'"

Quantitative analysis: ***Is there any attempt to make this a quantitative assay?*** **No.**

Dr. Stephen Little, CEO of DxS Ltd, a U.K. firm working with Amgen and ImClone on diagnostic assays, said, "We have not attempted to make the test quantitative…It seems a low or high level appears to qualify with response to these drugs. In this particular setting, most patients don't respond, so it might be better if we knew patients were 90% positive."

PUBLIC WITNESSES

Robert Erwin, C3: Colorectal Cancer Coalition. "Cancer drugs don't work very well or very long for many patients. Many treated patients get no benefit, and all patients are harmed in some way. For patients making a choice, it comes to the risk of harm with benefit vs. the risk of harm with no possible benefit. In K-ras we are dealing with a marker of no benefit, and that is an important consideration…We believe good practical judgment suggests that (the assay) can legitimately be used to guide the use of EGFR inhibitors."

Carlea Bauman, C3: Colorectal Cancer Coalition. She raised several questions about K-ras testing, including:

- What do we know about the biomarker?
- Are the results consistent with known method of action?
- What do we know about the assay? Is it new technology? Subjective or objective?
- Does it require black box calculations like (Genomic Health's breast cancer test) Onco*type* DX?

She concluded: "On K-ras, we strongly urge ODAC to recommend a label modification to allow some mention of these findings in the labels."

GlobeImmune official. "We proactively target K-ras in the development of our ras-targeted vaccine therapy. We probably are just beginning to understand the complex nature of (ras)… Perhaps it is over-simplified to say someone is either wild-type or mutant…and we should more fully characterize the mutations…In our hands, we still consider the gold standard to be 'double-stranded sequencing.' Commercial assays are a significant advance in convenience and ease...We think ras holds promise for predicting outcomes in cancer therapy, but there are implications of false positives – withholding potentially beneficial therapies. And the implication of a false negative means administration of therapy would be unlikely to confer benefit. How low is low enough for false negative? Is 5% enough? This is a very important question."

PANEL DISCUSSION OF FDA TOPICS

The FDA posed five discussion topics for the panel, with two introductory notes:

1. Assuming: Appropriate tumor sample acquisition and handling procedures were used, the assay for the biomarker has acceptable analytical validation, and clinical data would be obtained from randomized, controlled clinical trials.

2. Considering: Studies which met the pre-specified primary study endpoints and would not be intended as a mechanism to salvage failed trials.

TOPIC 1: When would it be appropriate to limit use of a drug to a subgroup based on retrospective analysis of one or more studies that were not designed to examine this subgroup? YES, on a case-by-case basis.

Please discuss the factors to be considered, including:

- Claims to be made – efficacy vs. safety (differences in risk:benefit) for the drug.
- Claims to be made for effectiveness and safety of the companion diagnostic test.
- Number of studies (replication of finding).
- The proportion of the intent-to-treat population in which biomarker results are available. What fraction of missing biomarker data in this entire population would preclude a decision regarding effects in a subgroup?

Ralph D'Agostino PhD, chair of the Mathematics and Statistics Department at Boston University, pointed to four general categories that, if all were met, might allow a believable analysis:

1. **The analysis should be hypothesis-driven** even though it is retrospective. "There should be validation built in… There is discussion of PFS vs. OS…but efficacy is very much driven by the original study."

2. **Where did the samples come from?** "I'm concerned that no matter how well you plan the study, if all you have is convenient samples, then you are in trouble…In some of the studies there isn't rigorous oversight of how samples are taken and stored…The idea of solid samples available is really going to drive this."

3. **Statistical power and multiplicity control** should be demonstrated with enough patients, enough data, enough samples, with multiplicity built in.

4. **Consistency and sensitivity analysis** should be built in. "Do we find consistency with interaction tests…Are the data sensitive enough that being positive or negative on the biomarker will make a difference?…And will the sample we are looking at be able to reproduce the original sample results so that this isn't a unique sample."

Other panel member comments included:

- *Dr. Gary Lyman, director of Health Services and Outcomes Research Program – Oncology at Duke:* "If the original endpoint is not reached in a prospective trial, the desire to go back and look at subgroups retroactively just doesn't really hold, and I think we should require two prospective studies where patients are stratified *a priori* based on the assay or the treatment was limited to a specific subgroup of the assay…It is extremely important that analyses be adequately powered within the subgroups, especially if the marker is more prognostic than predictive."
- *Dr. Harrington:* "There is a difference between regulatory approval and the march of science…I want to be sure we don't send a message as a committee that we can't learn anything from a trial that didn't meet its primary endpoint…It may mean additional trials for regulatory approval but not to imply that these trials are not useful …The issue in ascertainment is not how many (samples) you have but how you got them. Acknowledging you can't get them all highlights the importance of how you get them."
- *Dr. Simon:* "We've had conventional wisdom that you never trust a subset analysis unless the overall results are positive, and that protected us against data dredging, but that is an irrational rule of thumb now…We don't need that to protect us against data dredging. We need to distinguish between data dredging and the K-ras situation ...If we use the rule of never looking at a trial unless the primary endpoint is significant, that leads to erroneous conclusions. That rule of thumb needs to be given up."
- *Dr. Derek Raghavan, an oncologist from the Cleveland Clinic:* "We want to remember the comment of the patient advocate which essentially was a plea for common sense …The reality is we want to be careful we don't box (FDA officials) in a little corner where we set the bar so high that they can't make sensible decisions…One of the advantages of ODAC is that we don't have lawyers on the panel…Two companies came here to try to create a situation where they sell less product. That is unique… I've been a little uneasy that we have been raising the bar."

TOPIC 2: When would a prospective study, designed for the purpose of examining treatment effects on a pre-specified subgroup, be needed to establish treatment effects in this group? Probably not necessary if there are good retrospective/prospective trials.

- *Dr. Harrington:* "I think that my default would be if we have two very good retrospective/prospective trials that meet all the conditions, and they are the universe, then I don't think we need a prospective trial, but absent any of those conditions, that's when we need it (a prospective trial)."
- *Dr. Ronald Richardson, an oncologist from the Mayo Clinic:* "I think we need to spend more time being sure we have the proper markers…In the data presented earlier, if <20% of patients actually have objective response in the most favorable groups, the situation is very complex and requires further studies with the trials looking at the proper markers…I was struck with the repeated assertion that a randomized study of some of these drugs in WT K-ras CRC patients can't be done… One would wonder whether that conclusion that the studies can't be done is true."
- *The FDA's Dr. O'Neill:* "Do you have any advice on the disease to be sure you don't have a biased sample?" Dr. Simon responded, "You want to know all the details on the cases…whether there is institutional variability in treatment assignments and as much as possible about the issues on who you have ascertainment on and who you don't, so you can try to assure there is not a treatment difference on ascertainment…and also to try to understand what potential issues might be there in terms of generalizability."
- *Dr. Wyndham Wilson, chief of the Lymphoma Therapeutics Section, Metabolism Branch, National Cancer Institute:* "I think you need multiple studies showing the same thing…As someone who does a lot of biomarker work, the biggest block isn't the patient, it is the treating doctor. Again and again, I keep hearing it is the patients, but the enemy is us. That's where the roadblock is, I believe…My heart goes out to the patient who has been told he or she is K-ras mutant…In many cases they had to give up insurance to get in these trials…If you have been tested for K-ras, then your BlueCross (insurance) is not paying."
- *Virginia Mason RN, executive director of the Inflammatory Breast Cancer Research Foundation and the panel's consumer representative:* "One problem we've had with patients about accessing tissue is that imbedded in the consent is they give up any rights to that tissue or the ability to direct it. I think patients are becoming much more savvy about it."

To what extent is it important that the effect is consistent across PFS but not OS?

- *Dr. Jean Grem, a hematologist/oncologist from the University of Nebraska Medical Center:* "Part of the difference in outcomes was how the studies were designed. Panitumumab allowed planned crossover…but doing that made PFS the primary endpoint. But the other study in Canada was targeted to survival as they didn't have access to Erbitux unless the patient was randomized to it (no off-label access)."
- *Dr. D'Agostino:* "I walked away from those studies uncomfortable. There wasn't consistency between PFS and OS…I don't think you can walk away from that question easily."

- *Dr. Raghavan:* "I wasn't too bothered by that."
- *Dr. Wilson:* "Do PFS and OS have a place in settings like this?…I agree that if you are targeting different endpoints, there are design issues that could preclude you seeing other endpoints, but I think – in the absence of a quality of life benefit with PFS – we need to move more and more away from surrogates and more toward the bottom line: Are you living longer and with a better quality of life?"
- *Dr. Ronald Przygodzki, associate director, Genomic Medicine, Office of Research and Development, Department of Veterans Affairs:* "The consistency of the PFS data and the response rate data to me were sort of overwhelming…I don't know what else you could ask for. It was sort of a slam dunk…In third-line CRC with crossover do you really expect to see an effect on survival?"

Is there a fraction of missing biomarker data in a retrospective study that would make you say you must do a prospective study?

- *Dr. Harrington:* "Not unless it was near zero…It is not just size but how they get there…The further away from 100% you get in terms of ascertainment, the more red flags go up…In the data (on K-ras), there was high ascertainment except in one trial…When this type of situation comes up, it is always the A team that presents its data. Down the road, FDA will be looking at people with 25% ascertainment…The more people not ascertained, the more risk there is something crooked going on."
- *FDA*'s *Dr. Pazdur:* "You put a lot of conditions on – if you believe randomization is preserved, if you believe there is no bias – but how do you determine that in the real world?"
- *Dr. Harrington:* "The simplest rule is to get all the tissue, and that removes the issues of something selective missing, but that is probably not practically possible… The way I do it is to look at every case and try to understand if there might be some selective effect there that is hidden even from the investigators."
- *Acting panel chair, Dr. Janice Dutcher of Montefiore Medical Center:* "It needs to get on the website about K-ras and explain why the tissue is so important…I really think the patient advocacy community understands the importance of this prospective effort in community hospitals as well as academic centers."
- *Dr. William Funkhouser, a pathologist from the University of North Carolina Hospitals:* "With big primaries (tumors) like lung and CRC, there is no excuse not to have enough material."

TOPIC 3: Discuss the properties of clinical studies, originally designed for non-selected populations, that would make such studies *unsuitable* for demonstrating efficacy in a biomarker subgroup. Discuss in your answer potential problems associated with the failure to perform stratified randomization based on biomarker status, failure to pre-specify statistical adjustments for multiplicity, and incomplete ascertainment of biomarker ("convenience sampling"). The FDA should issue guidelines.

- *Dr. Harrington:* "The (FDA's) Critical Pathway initiative is terrific and well laid out. What I'm hearing in the committee is that the Agency will need to develop a set of useful working guidelines where the Critical Pathway doesn't apply…What jumps out at me is the failure to pre-specify statistical adjustment of multiplicity…You should not treat these as a data-dredging exercise…but you should force people to say there is a pretty good biological story that generates a hypothesis, question, etc."
- *Dr. Raghavan:* "While there is controversy in the literature…that should be a red flag…The thing that hasn't been talked about very much is the concept of too many variables…Keep it at one marker and one agent, looking for interactions where you might hope to see them… Multiple labs would make it unacceptable to me…There is a difference between biological and clinical relevance and significance vs. statistical significance. You will often get p=0.003, but that may still be a biological effect that is not important. It comes back to using common sense."
- *Dr. Joanne Mortimer, an oncologist from City of Hope Comprehensive Cancer Center:* "The goal is to do the biomarker on the primary (tumor) and on the recurrence to avoid the same problem that occurred with Herceptin."
- *Dr. Przygodzki:* "To me one of the most important things in the analysis is that it should be focused on a single biomarker…If a sponsor plans such an analysis, they should have to have an analysis plan and let the FDA know before doing the assay…That way the FDA would know presumably how many such markers they potentially looked at."
- *FDA's Dr. O'Neill:* "I think that is a good idea but impractical…because people are searching, searching, searching…Maybe we could move to a space where people would commit sooner, but I don't know who could pull that off practically speaking."

TOPIC 4: When is it acceptable to limit future enrollment to a biomarker-selected subset of an actively accruing clinical trial based on external information (e.g., results of another study)? What would be the primary analysis population? Would the answer depend on the proportion of unselected patients (i.e., those enrolled prior to the study modification)? When a biomarker would prevent patient harm, it is appropriate to use it to limit enrollment in ongoing trials.

Dr. Simon commented: "It is the responsibility of the DSMB (data safety monitoring board) to weigh the information…And they are the right people to do it because their responsibility is to the patients, not the sponsors or the investigators…Anyone else might have conflicts…When that kind of thing happens, I can envision two kinds of trials that would be ongoing:

1. "One is a trial in which the **biomarker-defined subset was something known at the outset** of the trial and incorporated into the design of the trial. Maybe because people originally didn't think that test-negative patients would benefit, but that information is never for sure…So, it was decided to go ahead and include the test-negative patients…Then some other trials provide information…In that situation, restricting entry probably is not too disruptive in terms of analysis of the trial…but the trials have been designed and had a primary analysis plan that included that predictive biomarker – either targeting an adequate number of test positive/negative patients and handling multiplicity, etc…so that is not so disruptive.
2. "The more difficult situation is like with K-ras **where the information comes up**, and it is information that was not available at the start of the trial, so the trial was not started with that as the predictive biomarker. I'm afraid I don't have any great rules on how you deal with it…In some things, there are no rules. They have to be dealt with on an individual basis, with best judgment available. If something as important as external information leading you to a negative predictive biomarker came about, it would be somewhat ridiculous to ignore that in the analysis of the trial…So, I would think that probably for the trial you have to look at the effects overall and the test positives and negatives."

Dr. Grem added, "I thought the situations where it makes the most sense to stop a trial and exclude patients are in settings where you have information that a biomarker would predict a high risk of patient harm…That perhaps they were not able to metabolize or deactivate the agent they were given, so there is a high risk for toxicity. Or, in the situation where you have no chance of benefit. Those are pretty clear cut. The things that are more difficult are when we think patients may be *more likely* to benefit. That would be more difficult. I think it is easier if there is no chance vs. you ***may*** benefit…I think it is kind of a safety issue…where with no chance of benefit the risk becomes unacceptable…In all the other areas, like we think you might be more likely to benefit so let's not randomize those patients, I don't think that would be really good."

The FDA's Dr. Pazdur clarified what the FDA was asking the panel: "We agree the DSMB has the primary responsibility, but we wanted more granularity. For example, what was the endpoint that one used to make this decision from external information? Was it one trial or six trials? What was the effect of the endpoint? Was it 50% doubling in PFS, six-weeks improvement, or a 5% difference in response rate? And if it was done in a different disease setting, like a very refractory setting, should adjuvant or first-line trials be stopped? We wanted a conversation on that. There is a high degree of subjectivity that could come into play here – affect size, endpoint used, the constant issue of replications (how many trials), implications in other diseases or other disease settings in the same disease." Panel member responses to this included:

- *Dr. Grem:* "After going to ASCO, the (K-ras) trials that struck me (impressed me) were third-line in monotherapy vs. best supportive care. Those were pretty convincing and almost identical in that patients who had mutant K-ras had no benefit…whereas there was a pretty big response with wild-type K-ras…If a patient has mutant K-ras, it is unlikely they will regain normal K-ras…So, if the patient is mutant, I think they will always be mutant…We can argue about how many (mutant) cells, but that was pretty striking data…In light of that, I thought that for the CALGB study, it makes sense to go ahead, stop, modify it so only K-ras wild-type patients are randomized to Erbitux…What to do with the rest of the trial? I don't have a control, but I still think an increased sample size to look at benefit in K-ras wild-type patients…but I think they should look at overall effect in all patients and then do a secondary analysis in an expanded trial…The things I don't know are like in an adjuvant study, which is modified so patients with wild-type K-ras are eligible to FOLFOX or FOLFOX + Erbitux…I wonder if they shouldn't just have been followed."
- *Dr. Michael Link, chief of the Division of Hematology/ Oncology at Stanford:* "Patients read it (K-ras)…and wonder why they should participate in the trial. It is difficult enough for clinicians when the DSMB said keep randomizing patients, but the patients may not want to get randomized and could kill the trial no matter what the DSMB wants to do."
- *Dr. Wilson:* "While there may be less benefit in the adjuvant setting, it may not be zero…I guess I would have been very cautious to have thrown people off the adjuvant trial who had K-ras mutations…We also know that with other drugs, as you use them in more and more (sick) patients, they work less well."
- *Biostatistician Dr. Simon:* "I know guidelines are useful, but in many of these complicated situations, guidelines only carry you so far…I think there probably needs to be somewhat more recognition of the actual relevance of potentially changing analysis plans prior to the analysis of data, while it is still blinded, when the study has already started."
- *The FDA's Dr. O'Neill:* "(There are examples) where a lot of smart people changed the endpoint midstream, and at the endpoint, they lost. But if they hadn't changed it, they would have won…My concern is…(when) you are fooling around with one study, and you are making some mid-course changes where you don't even have a good analysis planned…We have a version of this in the results in other areas where the results either look better or worse

in or outside the U.S. This is a version of a subgroup analysis."

- *Dr. Funkhouser:* "If we stop a trial because we are convinced there is no benefit for some patients with a certain genotype…that sends a message that there is no hope… So, you need to be statistically right as well as emotionally confident. And just because a person is K-ras wild-type doesn't mean other proteins in the same pathway, like B-raf, aren't mutated."

TOPIC 5: Please discuss the importance of timing and rigor in determining the analytic performance of the companion diagnostic test. Ideally, diagnostic tests would be developed with a drug, but that may not be practical or realistic.

This was discussed first because of a panel member schedule issue. Dr. Przygodzki said, "If…everyone in the world starts using this, we have to have some guidelines on the typical approach to make this as even a type of test in the sense of accrual of tissue and that the test itself is relatively standardized. On top of that…what is the true cutoff for positivity? If we go back to current gold standard – bidirectional sequencing – we are looking at 20% mutations. With the current methods, the sample is much more refined, and one can identify a small percent of cells that are mutated. Is a 5% mutation entirely mutated? What if one uses 20% as the gold standard for mutation? We need a clear cutoff to establish what is and isn't mutated."

Are there any guidelines for sample collection? Dr. Przygodzki said, "There are general guidelines – about margins, on samples per centimeter, essentially all the lymph nodes (in the area). It isn't truly different from place to place. If you miss or under diagnose, it is criminal…so one goes the extra mile to make the diagnosis as accurate as possible." Dr. Wilson said, "I think we would all agree it is optimal to do this prospectively, but the reality is many targeted agents one finds out over time are not truly targeted to any biomarker and may go from mechanism-based biomarker to prognostic biomarker and may affect both patients with mutation or not…In the example at hand (K-ras), one of the pitfalls, besides statistical ones, in terms of doing the current study after the fact, is that among those patients EGFR negative by IHC, we don't know if wild-type K-ras might identify a patient who might benefit from a drug. So, do you think missed opportunities come about as well by doing this later? I think at the end of the day …this is going to be driven by the availability of these types of tests…and often, they are simply not going to be available initially during the upfront studies." Dr. Simon added, "Ideally, you want an analytically validated test used prospectively in pivotal trials, but because of the complexity and because science is often out of sync with clinical development, that often won't be the case."

What are predictive markers? Are there some too sensitive? Dr. Przygodzki said, "I don't think anyone knows what the cutoff is to make something positive or negative…If you get to the single-cell level, you may find there are 2-3-4 mutations going on. Does that actually mean there is truly a mutated tumor? I don't know. I don't think anyone knows." Dr. George Netto, a pathologist from Johns Hopkins Medical Institutions, added, "I think that is why it is crucial to use the test in the same method or cutoff as the pivotal trial. You cannot use a mutation that excludes you from treatment, and then adopt a fancier test…because you aren't using the same standard, the same cutoff, and same test." Dr. Raghavan said, "We are not, around this table, going to define the ideal test because it doesn't exist…You can talk about the importance of reproducibility, but now it doesn't work that way – temperatures vary, etc…I'm not suggesting a gigantic biorepository…but we need some frame of reference."

The FDA's Dr. Keegan suggested, "FDA could reach agreement with a sponsor to store the samples themselves under the postmarketing commitment, so samples are always available for the future…The heart of this question is…What kinds of things should we do, if there is a postmarketing commitment, that there be a plan for assessment of future technology? What kinds of things should we have on hand at approval or shortly thereafter, so we can (follow) these commitments up more intelligently?" A panel member said, "Perhaps all Phase III licensing trials for novel targeted agents should have mandated prospective tissue acquisition with a provision made for storage by the sponsor." Nicole Vesely, PharmD, an FDA staff member, said, "I think it will really be based on the type of test you are actually going to be doing. I think K-ras looks like it is a relatively early event, so it should be present in most patients upfront. But p53 is stable in large cells and unstable in CLL (chronic lymphocytic leukemia), so it is a moving target." Dr. Przygodzki added, "On the mandate for tissue: It is a great idea scientifically to mandate them prospectively, but I don't think our IRB (institutional review board) would approve that."

ASCO RECOMMENDS K-RAS TESTING

In January 2009, the American Society of Clinical Oncology (ASCO) issued a "preliminary clinical opinion": "All patients with metastatic colorectal carcinoma who are candidates for anti-EGFR monoclonal antibody therapy should have their tumors tested for K-ras mutations…If K-ras mutation…is detected, then patients…should not receive anti-EGFR monoclonal antibody therapy as part of their treatment."

In a press briefing in advance of the ASCO Gastrointestinal Cancers Symposium, Dr. Veena Shankaran of Northwestern University said that determining a patient's mutation status would lead to better targeting of Erbitux to the patients most likely to benefit from it. She estimated that mCRC treatment costs could be reduced by as much as $604 million a year if all patients were tested for K-ras mutation status, and the savings might be even higher if associated costs (e.g., clinic visit, infusion time, toxicity management) were included. ♦