Author's response to reviews

Title: Evaluation of a novel rash scale and a serum proteomic predictor in a randomized phase II trial of sequential or concurrent cetuximab and pemetrexed in previously treated non-small cell lung cancer

Authors:

Michael L Maitland (mmaitlan@medicine.bsd.uchicago.edu)
Matthew R Levine (mlevine1@medicine.bsd.uchicago.edu)
Mario E Lacouture (lacoutum@mskcc.org)
Kristen E Wroblewski (kroblewski@health.bsd.uchicago.edu)
Christine H Chung (cchung11@jhmi.edu)
Ilyssa O Gordon (Ilyssa.Gordon@uchospitals.edu)
Livia Szeto (lszeto@medicine.bsd.uchicago.edu)
Gail Ratko (gratko@sbcglobal.net)
Keyoumars Soltani (ksoltani@medicine.bsd.uchicago.edu)
Mark F Kozloff (mkozloff@bsd.uchicago.edu)
Philip C Hoffman (phoffman@medicine.bsd.uchicago.edu)
Ravi Salgia (rsalgia@medicine.bsd.uchicago.edu)
David P Carbone (d.carbone@vanderbilt.edu)
Theodore G Karrison (karr@uchicago.edu)
Everett E Vokes (evokes@medicine.bsd.uchicago.edu)

Version: 4 Date: 20 August 2013

Author's response to reviews: see over
Dear Dr. Tan:

We are grateful for the opportunity to address the reviewers’ thoughtful comments. Attached are two versions of the manuscript, one clean, and one with the changes tracked in MS Word. Below is the requested point-by-point response to the reviewer comments.

Referee 1 Comments
This is already a completed clinical trial and any comments are discretionary. This main thrust of the study attempts to analyze a novel proteomic marker for response to cetuximab. Cetuximab has since been shown to have limited role for NSCLC, and so this data may no longer be relevant. Moreover the proteomic marker appears to mark for disease stability or non-progression rather than response (figure 3) limiting its usefulness in its application.

We agree with this reviewer’s concern that this study has little value in directly informing current standards of care in lung cancer. This is common for phase II combination therapy trials in all of oncology. Our purpose was to analyze our rash and serum proteomics data in a fashion that would support inferences on their relative performance in this trial to guide their further development. We aim to help others expend further effort only on the biomarkers that show early evidence of possibly having greater impact in a treatment setting other than second-line non-small cell lung cancer. We have modified the text throughout to make this point clearer.

Referee 2 Comments

1. Their original statistical design called for 40 patients per arm and due to practice changes calling for study termination, this study needed to be terminated after a little more than half accrued, therefore their observations mostly seem to be in the thought-provoking range and there are concerns about false positive and negative observations. That being said, publication of the results appears to be of value and one suggestion might be just appropriately softening their statements as to the impact of their observations.

We appreciate Referee 2’s many thoughtful comments and concerns. We have inserted language in the abstract and throughout the text to “soften our statements and the impacts of our observations” as recommended:
- Background of the Abstract: Inserted “In a phase II trial that was closed to accrual because of changes in clinical practice we examined the relationships among candidate biomarkers, quantitative changes in tumor size, progression-free and overall survival.”
- Discussion of the Abstract: Inserted “were associated with” NSCLC outcomes. “Although in a small study, these observations were consistent with results from larger retrospective analyses.”
- additional modifications in response to specific points are described below and tracked in the new version of the manuscript.

2. The non-conventional imaging endpoints used for this study are not well explained and should be more clearly stated for the readership. Ideally, RECIST measurements and correlation of two imaging assessments should be added.

- In the fourth paragraph, second sentence we changed the sentence to read “A component of the trial was to conduct preliminary development of biomarkers for cetuximab.”
We have inserted two additional explanatory paragraphs at the end of the Background on pages 3-4. “The trial closed prior to achieving its initially intended accrual goals. The original hypothesis to be tested by the trial was to determine in patients with NSCLC refractory to previous chemotherapy whether concomitant treatment with cetuximab and pemetrexed improved progression-free survival compared with cetuximab monotherapy. We estimated a sample of 40 patients per treatment group based on many assumptions, including a relatively modest difference in outcomes between the study arms. Our estimates were inaccurate, and patients who were randomized to the combination therapy arm had much better outcomes than patients in the monotherapy arm. Additionally, large clinical trials of pemetrexed revealed advantages to early initiation of pemetrexed after first-line therapy, and single-arm studies of cetuximab in unselected patients revealed no benefit to cetuximab monotherapy in this population. Therefore the study was closed after reaching half of its intended accrual. The preponderance of available evidence now supports that the administration of pemetrexed in the second-line treatment setting would be superior to cetuximab monotherapy. We report the results of this trial not to demonstrate a definitive assessment of the difference in the randomized treatment arms, but rather to provide the foundation for this biomarker development exercise.

The trial was intended to study pharmacodynamic biomarkers in unselected patients and collection of tumor tissue was not incorporated into the study design. But the study called for all patients to provide pre-treatment serum samples and to undergo standardized, prospective rash evaluations. There are many limitations to the successful development of valid clinically relevant markers concurrently with novel anticancer agents. The need to collect material for markers prospectively, but the observation time required to accumulate outcome data with which to qualify biomarkers makes the process slow and expensive. In this manuscript we have examined the data from this trial with 3 goals: 1) to explore the use of quantitative changes in tumor burden at an early time-point in a trial as a measure of drug effect (rather than RECIST-based response or time to progression), 2) to determine whether the prospectively collected biomarker data in this trial are consistent with retrospective analyses of larger clinical trials, and 3) to perform a preliminary estimate of the value of intensive, quantitative rash assessments as a candidate biomarker for future studies of EGFR inhibitors.”

In the Methods section on page 6 we modified the text to clarify our strategy with the following: “The primary motivations for use of Response Evaluation Criteria in Solid Tumors (RECIST) in phase II clinical trials and the shortcomings of response rate and progression-free survival as endpoints in small randomized trials have been well addressed elsewhere[19, 23-28]. To maximize our sensitivity for detecting differences in performance of biomarkers for cetuximab, we applied the novel quantitative variable derived from the modeling studies, the log ratio of tumor size at 8 weeks of treatment, as an experimental measure of treatment effect in the context of these data.”

In the case of this trial, making estimates and inferences from this incompletely validated endpoint strategy is our only opportunity to extract value from this clinical trial. We know from the cited modeling studies that the variance in RECIST measures of response rate and progression-free survival will be too variable to expect any correlations of significance and this is the reason for displaying the data as we have in Figures 1, 3, and 4. Because of measurement imprecision for detecting treatment effects by “response” and progression-free survival, our small sample size, and our imperfect randomization, we cannot demonstrate any statistically significant correlations. Nevertheless, we might inform future development of biomarkers for EGFR inhibition by examining our multidimensional data (treatment by log ratio tumor size by rash rating score change by proteomic predictor) to infer where further
efforts might be best devoted. Even with all of the shortcomings of this paper, with the run-in period, our rash stratification worked well. But for all of our efforts in developing this quantitative rating scale, it worked no better than categorizing patients as “rash” vs. “no-rash” which is what was found in Gatzmeier, et al. with much less intensive assessment of rash than we had prospectively incorporated here. Similarly, our proteomic marker predictions were consistent with the outcome observations. This is not definitive work but suggests either a rash/no-rash categorization or a serum proteomic marker strategy are worth testing further in cetuximab studies, especially in any additional studies of cetuximab added to front-line chemotherapy.

3. There seemed to be 12 drop-outs, the authors should explain why they faced such a large drop-out rate.

- We have inserted text into the Results section under Patient Characteristics on page 6:
“...The drop-out rate was not unusual for a second-line therapy trial in the pre-pemetrexed era, especially because this trial pre-specified that only patients who completed the 3-dose, 2-week run-in of cetuximab would be considered evaluable. Five patients were withdrawn because of infusion reaction to cetuximab, 3 patients withdrew because of inconvenience of commuting to the clinic site, 1 patient could not get pain adequately controlled and withdrew to begin immediate cytotoxic therapy, 1 patient had symptomatic progression of bony metastasis at day 15, 1 patient died suddenly and unexpectedly at week 4, and 1 patient withdrew prior to the first imaging evaluation for intolerability of grade 2 rash...."

4. The outcome of patients on the cetuximab arm overall seemed extremely poor with an overall survival of 3.5 months. One wonders how much this was the result of some introduced bias as a result of potentially too sick patients to join as better results would be expected even with best supportive care. The authors should comment on this more. Certainly one significant bias seems to be highlighted by the very large percentage of poor Veristrat group patients in arm A and indeed it needs to be highlighted that this bias makes it impossible to know whether the differences between the arms are treatment-related or related to accidentally biased assignment of patients to the two arms and this should be highlighted more in Discussion.

This is an important point, although all evaluable patients met the study entry criteria, competing trials run at our institution concurrently introduced a bias for accrual to this trial toward more ill patients than usual. To make this problem identified by the reviewer clearer we entered the following text to the Discussion on page 8:
“As a small trial, not meeting its intended accrual goals, there are clear limitations to this study. The setting of advanced NSCLC meant many patients deteriorated during or immediately after investigational treatment. As the value of information was not clear at the time of conducting the trial, we did not assess candidate molecular markers on patients’ tumors and we did not stratify the randomization by tumor histology. We had competing trials at that time that required information on tissue type for enrollment and a PS of only 0 or 1. This study unintentionally enrolled a typically more ill population with few never smokers (10%), fewer than usual women (under 40%), and many patients with non-adenocarcinoma or unknown histology (70%). The actual sample size meant only large effects, such as the benefit of immediate initiation of pemetrexed, could be detected with conventional measures such as progression-free and overall survival. The randomization did not effectively free the assessment from bias. This is highlighted by the unusually poor median-survival time in Arm A, 3.5 months is even low for a best-supportive care only trial. Arm A had fewer women, had patients with a higher median age, and more patients with serum proteome-profile-predicted “poor” outcomes with cetuximab monotherapy. It is therefore unclear the extent to which the treatment assignment versus the small sample size have biased the outcomes of this trial by study arm.”
5. As standard premedications for pemetrexed include Decadron, one wonders whether this might have influenced rash rates on arm B and rash scaling on cycle 1, day 8. The authors should comment on this possibility.

As depicted in Figure 4, the change in rash by C1D8 was actually greater in the Arm B subjects than Arm A subjects and we inferred that at this one timepoint the dexamethasone administration did not affect the severity of rash.

6. As pemetrexed is felt to have poor activity in squamous cell patients, the authors should comment on impact of histology on outcome measures

We inserted the following text on pages 8 and 9 in addition the insertion above. “Cetuximab monotherapy has no evident value in the treatment in unselected NSCLC patients in the second-line setting. Some could argue that cetuximab therefore has limited if any role in combination therapy. However, the data from Gatzemeier, et al[7]. suggest that a biomarker-based selection of patients who should receive cetuximab added to standard chemotherapy could yield improved outcomes over those reported for the cetuximab arm in the FLEX trial. This circumstance is increasingly common in oncology therapeutics, an agent that has limited but evident benefit in combination cannot be used ethically as monotherapy. Therefore, the early development monotherapy trials become an important opportunity with which to characterize candidate biomarkers and conduct preliminary validation and comparative estimate studies. In this “pure” setting, investigators can determine the typical time course, intensity, and interindividual variance in candidate markers. In a single disease, investigators can conduct preliminary comparisons to make estimates regarding which markers represent the best opportunities for future validation and qualification studies in large trials, including combination therapy trials.”

7. The authors find that the Veristrat good predictor class correlated with favorable changes in tumor size- was that noted for both treatment arm or arms combined?

Due to our small sample size, and missing data on 10 of the 43 subjects we did not undertake this analysis. We simply display all of the data in Figure 4 for the reader to review.

8. The overall data for arm A clearly highlight that cetuximab has no discernible single agent activity in this setting (and would not have made it into phase 2 of a Simon two-phase design). If the authors agree with this statement, this should be stated more firmly in the Discussion. In essence, this aspect should also lead to some consideration for the actual need of developing biomarkers for cetuximab-based therapy and this should be more fairly stated in the Discussion on aspects of future development plans.

We think the insertions in response to points 4 and 6 above also address this concern.

9. Many patients, especially on arm A were diagnosed with progressive disease/CNS metastases on brain MRI- were these symptom-directed MRIs. It seems that baseline brain MRIs were not needed. Would results be substantially different in arm A as to survival data if the early brain progressors were excluded?

- We entered into the Methods section on page 5 the clarifying statement
  “Baseline brain MRimaging was not performed. Patients who developed central nervous system symptoms were referred for imaging.”

- Overall sample size was too small to objectively address this good question.

10. Figure 4 is very difficult to understand and the authors should explain this figure better in the figure legend and make clearer statements as to what conclusions they reach by it
We have generated a more complete Figure legend to address this concern.

“Figure 4. Change in tumor size vs. change in EIR rash score. This is a plot of each individual, evaluable patient from this trial. The x-axis represents the change in EIR from baseline to C1D8. The y-axis represents the change in tumor size by log ratio of tumor burden at week 6 to baseline. Each patient is further represented by which arm to which they were randomized (red for Arm A, blue for Arm B) and serum predictor (x for poor, circle for good). The plot reveals potentially informative trends: 1) the top 5 tumor responses (the most negative log ratios) were among subjects randomized to Arm B, all of whom had “good” predictor status, but had rash changes on the EIR of 0, 1, 2, 4, and 5; 2) Of the early progressors (11 subjects with log ratio > 0.2 at or before the first imaging session), 8 had some evidence of rash and 5 had among the most severe rashes with an EIR rating change of 4 or 5; 3) No patient with a “poor” serum proteome predictor had any tumor shrinkage at all, but only 2 of these subjects were randomized to receive pemetrexed at the end of the 2-week run-in. “