May 1, 2011

IOM Committee on Genomics

Dear Colleagues of the IOM "omics" Committee:

At your request, I have prepared a statement describing my involvement in the recent controversy regarding a genomic signature to predict outcomes in lung cancer, developed at Duke University. My laboratory has been involved with the gene expression profiling of lung cancer for the last decade and was part of the original Director's Challenge Program funded by the NCI.

In 2002, we published one of the first papers describing the preliminary discovery of a prognostic gene signature in non-small cell lung cancers collected at the University of Michigan. Since there were several groups investigating potential genomic signatures in lung cancer in the Director's Challenge, I worked with Jim Jacobson of the NCI (deceased) to organize an effort to conduct a large scale, collaborative consortium study of lung adenocarcinoma. We proposed to establish a multi-institutional tissue and data-set containing over 400 tumors that would serve as a resource for future research. Our proposal included using this dataset to train and test individual prognostic gene signatures and directly compare their performance. Tumors from the University of Michigan, Memorial Sloan-Kettering Cancer Center, Ontario Cancer Institute/Dana Farber, H. Lee Moffitt Cancer Center were included, as were specimens from patients who participated in a lung cancer clinical trial within the Cancer and Leukemia Group B (CALGB), protocol number C9761.

Several guiding principles of this Consortium were reached by consensus. The first was the agreement that all data would be shared and made publically available, but that none of the data would be published until a main paper, describing the overall dataset and the genetic profiles we obtained, were first published as a whole. In this regard, after several years of collaborative work, we published the main manuscript in Nature Medicine in 2008, describing discovery and validation of a prognostic signature in a training set cohort and two validation sets, respectively.

The second principle was incorporated into a signed document via the NCI-funded collaborations that any manuscripts from the data had to be provided to collaborators at least 60 days before publication.

My laboratory was responsible for the specimen processing (sectioning, etc), isolation of mRNA and performance of gene expression array assays in order to develop a training set from the specimens submitted by the Consortium. The data were supplied to Kerby Sheddon, PhD, a bioinformatics expert here at the University of Michigan, for both storage and bioinformatic analyses.

The appearance of part of this dataset in a manuscript by Potti *et al* in the New England Journal of Medicine in 2006 has been problematic since its publication. At your request, I am providing my concerns and comments regarding the Potti *et al*, NEJM 2006 paper as it relates to the IOM committee investigation. In the recent Cancer Letter I also provide most of this information but the main points are as follows.

A. In 2005 (I did not keep notes at that time and I am not sure of the specific date, but clearly far preceding the publication of the main Consortium manuscript in 2008), I was asked by Dr. Potti for access to the data from the consortium regarding unpublished mRNA dataset on lung adenocarcinomas. This conversation took place via teleconference, in which he was in the office of Dr. Matthew Meyerson at Harvard and I was in my own office in Ann Arbor. I told him that, according to our consensus agreement, the data could not be available until we made them public after the main manuscript was published.

   Dr. Meyerson has subsequently stated to me that he doesn't recall this conversation. Therefore, although I am certain it occurred, I cannot document it and I am not sure of the exact date.

B. One of our collaborators on the Consortium project, Dr. Robert Kratzke at the University of Minnesota was one of the study chairs for CALGB 9761. Prior to publication of our main manuscript, he requested the subset of gene expression data that we profiled from the ~100 CALGB tumors. Since his request involved a cooperative group trial (as opposed to the single institution specimen data for the rest of the consortium), I agreed to provide the gene expression array data to him and his colleagues. I was not made aware of the exact use of the data, and there was no discussion that the data were to be provided to the Duke group, who, as noted, had asked for it previously.

C. Technically, the CALGB samples were problematic, as many had been potentially partially thawed and appeared to be not useable. To supplement these specimens in order to have sufficient power to accomplish the profiling objectives of the CALGB, tumors from the University of Minnesota tumor bank, collected and maintained by Dr. Michael Madeus, a colleague of Dr. Kratzke, were added to those from the CALGB study that we were able to profile. Indeed, my institution, via NCI, provided $10,000 to the University of Minnesota to support retrieval and mailing to us, in addition to the resources required to process and analyze these specimens, in order to maintain the collaboration.

D. In spite of our agreement not to publish separately until the main paper was released, and in spite of my not agreeing to Dr. Potti's request, these data were incorporated by the Minnesota and Duke groups to validate the Duke prognostic metagene signature, resulting in Figure 4B of the Potti *et al* NEJM 2006 paper. In this paper, the authors generated a signature from a training cohort of 89 patients collected at Duke, and then applied the signature to specimens from two validations cohorts: 25 patients treated within the American College of Surgeons Oncology Group (ACoSOG) and 84 patients described as being "from the prospective Cancer and Leukemia Group B (CALGB) 9761 trial".

   As far as I can tell, the specimens used for Figure 4B were a combination of those collected from patients who were enrolled in CALGB 9761 and from patients treated at the University of Minnesota but not enrolled in CALGB 9761.

Regardless, the partial dataset derived from the entire consortium collaboration, which had not yet been published, was made publically available (public posting of the supporting data is NEJM policy). Moreover, no other investigators, including myself, were extended the agreed-to courtesy of reviewing the manuscript with a 60 day grace period.

E. Perhaps more importantly, inclusion of this data set in the Potti *et al* NEJM 2006 raised major scientific and methodologic concerns in addition to the political disagreements outlined above:

1. In Methods, Potti *et al* stated that the expression data were generated using Affymetrix 133 Plus 2 array chips. This statement is incorrect. The array data were generated in my laboratory at University of Michigan, and we consistently used Affymetrix 133A array chips.

2. Potti *et al* stated that all the tumors were from CALGB 9761, yet as noted, a significant percentage were actually from the University of Minnesota supplement dataset, and were never given a CALGB number.

   They said that the CALGB samples without numbers were “CALGB run-in samples” although they were not all CALGB 9761 tumors. We do not know if this is true, but this is not stated in the manuscript anywhere.

   Tumor sample annotation in Supplementary Material demonstrates that the samples in question do not have CALGB numbers, and in other areas the annotation is incorrect in many cases. After this concern was raised, they subsequently corrected these errors, but only partially.

3. In Figure 4 of the NEJM 2006 paper, Potti, *et al* used a set of 25 lung cancer specimens collected by ACoSOG (Figure 4A) and 84 specimens that they state were from patients who participated in CALGB 9761 (Figure 4B) to validate this signature (again, in spite of my explicit request that they not do so at that time).

   Potti *et al* state in the legend for Figure 1 (the CONSORT diagram) that 7 of 91 of the CALGB tumors were “excluded from analyses on the basis of inadequate quality of the messenger RNA”. This statement is frankly untrue. It would have been impossible for the authors to have excluded any specimens based on RNA quality, since all RNA was harvested and examined for quality by Agilent bioanalyzer at University of Michigan prior to RNA expression array analyses, and I personally examined each bioanalyzer profile. Any specimens with inadequate or poor quality RNA were excluded by me prior to running the arrays or posting any of the data, and therefore the RNA quality of all of the more than 100 specimens to which they had access was of highest quality. I have maintained the RNA bioanalyzer profiles in my office, and I have no recollection of having shared them with any of the authors.

Thus, they only had access to high quality expression array data, and they could not have excluded any specimen based on RNA quality concerns, since they never had these data.

It is conceivable that they, or their collaborators in the CALGB statistical center to whom they state they submitted the predictive results for associations with the clinical outcomes, had legitimate reasons for exclusion of the specimens, but the reason stated in the figure legend for Figure 1 are not substantiated.

Following publication of the Potti *et al* paper in the *NEJM* in 2006, Jim Jacobson and I notified the NEJM of these many inconsistencies. We requested that the paper be retracted and/or corrected. However, the NEJM editor referred us to the authors, and when contacted, Dr. Nevins refused to take any appropriate actions. As a compromise, both the authors and the NEJM editors agreed that the array data would only be available on the NCI website, and that it would be removed from those maintained by Duke and the *Journal*.

In summary, the Duke metagene signature model appeared from the start to separate high and low risk groups with astounding magnitude, in spite of the well-documented heterogeneity of adenocarcinoma of the lung. The reports by Potti *et al* raised considerable enthusiasm in the research community. However, after the Consortium completed our analyses (published in Nature Medicine in 2008), none of the statisticians involved with our study were able to replicate the Duke result in the Consortium dataset, primarily because key information was not provided to allow successful replication.

This episode raises serious concerns regarding both the integrity of the investigators, but perhaps more relevant to the deliberations of the Committee, the response of the *Journal* to the many inconsistencies that were raised, both by us and others in the research community, and the lack of appropriate outlet for us to raise our concerns

Thank you for the opportunity to comment. Please feel free to contact me if I can be of any further assistance.

May 9, 2011

David Beer, Ph.D.