15-Month Progress Report

QIBA Project #3a

QIBA Committee/Subgroup: Volumetric CT Technical Committee

Project Title: Validation of Volumetric CT as a Biomarker for Predicting Patient Survival

PI: Binsheng Zhao, D.Sc.

Institution: Columbia University Medical Center

Project Overview: The goal of this project is to validate volumetric CT as a better imaging biomarker for predicting patient survival using a retrospective dataset from a large, multicenter Phase II/III clinical trial in advanced colorectal cancer treated with a targeted therapy. About 560 patients enrolled to this clinical trial had measureable disease per RECIST definition and will be included and analyzed in this study. Our in-house lesion segmentation algorithms developed for solid tumors will be applied to assist radiologists in obtaining tumor volumes. The expected outcomes of this project include (1) (**Aim #1**) obtaining the knowledge about the levels of intra-reader and inter-reader variability in measuring total tumor volume/burden and change in tumor volume/burden (uni-dimensional and bi-dimensional measurements as well), and (2) (**Aim #2**) evaluating the value of the volumetric CT in predicting patient survival.

Project Update (Aim #1):

From May through July 2012, we added a second radiologist's repeat measurement, performed a thorough quality assurance test on all of the measurement results (incl. Aims 2's) and re-analyzed the updated data using Bland-Altman method. To date, this analysis has been completed. Part of our preliminary results was reported at ASCO 2012 (abstract appended).

Project Update (Aim #2)

To date, we have completed computer-aided tumor volume measurements (uni- and bi- as well) on all available CT scan time-points (4 - 5 scans per patient, on average) in all of the 560 patients. After we had locked our longitudinal measurements, we received the patient survival etc information. Out of 560 patients, only about 250 patients' survival data are available. Our measurement data has been merged with the patient clinical outcome data (e.g., survival) and was sent to our statistician for analysis.

ASCO 2012 Poster

Relationship of variability in tumor measurement and response

B Zhao, S Lee, HJ Lee, J Qi, T Persigehl, Y Tan and LH Schwartz

Background: RECIST is widely used to evaluate anti-cancer therapy efficacy. This study explored variability in reporting tumor change and response to therapy due to both target lesion selection and measurement.

Methods: 2256 measurements were performed in CT scans of chest, abdomen and pelvis from 30 patients retrospectively taken from a multicenter Phase II/III colorectal clinical trial. Using RECIST, three radiologists interpreted baseline, 6-wk and 12-wk scans in the following manner: (1) Radiologists independently selected and measured target lesions, (2) one radiologist's target lesions were re-measured by the other two and (3) one radiologist re-measured the same scans in the above manner at an interval of greater than a month to prevent memory recall. Measurement variability in total tumor burden (TTB) on relative changes at 6-wk and 12-wk from baseline was analyzed for inter- and intra-reader target lesion selection, inter- and intra-reader measurement using Bland-Altman method, and agreements on RECIST categorical responses were assessed by kappa coefficient.

Results: When the same target lesions were used, inter- and intra-reader variability in TTB on relative changes at 6-wk and 12-wk from baseline were similar; all had 95% limits of agreement within (-15%, 15%). Kappa coefficients for RECIST were 0.74 (6-wk) and 0.87 (12-wk) for inter-reader and 0.64 (6-wk) and 0.88 (12-wk) for intra-reader to report responses. When radiologists independently selected and measured target lesions, variability in relative changes was within (-17%, 16%) at 6-wk and (-24%, 23%) at 12-wk for inter-reader and (-16%, 16%) at 6-wk and (-14%, 18%) at 12-wk for intra-reader interpretations. Kappa coefficients were 0.66 (6-wk) and 0.75 (12-wk) for inter-reader and 0.63 (6-wk) and 0.80 (12-wk) for intra-reader to report responses.

Conclusions: Differences exist in measuring tumor change. The magnitude of change in categorical RECIST response is quantifiable. The largest differences are when radiologists independently select and measure target lesions, the smallest when one radiologist repeats measurements on identical target lesions. The variability may impact the reporting of categorical responses and trial results, especially in a single arm study.