QIBA/NIBIB Round-2 Project #15a

6-month Interim Report (25JUL2012)

Project Title: Extension of Assessing Measurement Variability of Lung Lesions in Patient Data Sets: Variability under Clinical Workflow Conditions

PI: Michael McNitt-Gray, PhD

Updated (except milestones) July 3, 2012

Hypothesis

That the minimal detectable change in tumor size - using measured tumor volumes made by radiologists on thin section CT images - will be smaller when using a side by side ("more clinical") review setting than when using an independent review setting.

Project Description

The purpose of this project is to *extend* the data collection and statistical analysis of the QIBA Volumetric CT committee's 1B experiment, which is investigating the minimum detectable change in lesion size from patient datasets imaged on CT. That project used: (a) "Coffee Break" CT image datasets from 32 NSCLC patients who were imaged twice over a short (15 minute) interval on the same scanner using thin (1.25 mm) slices; (b) one lesion was identified for each patient, (c) Image data was marked up by five radiologists at RadPharm (now CoreLabs); (d) each reader marked the lesions on each of the repeat scans to obtain measures of volume, single longest diameter and bi-dimensional diameters. This data was previously collected and initial analyses have been performed.

In that previous project, each reader performed the image markup as an independent reading, with no access to the results or images of their markup of any previous session. This was done because the only data used in that study were limited to "Coffee Break" experiment cases. If readers were allowed to both: (a) know that these were cases with no change and (b) see their markings on one scan while making markings on the second scan, then those results would be considered very biased as readers would know what the answer should be (they would inherently know that the markings should match and that change should be zero). Using that independent reading paradigm, the initial results indicated that there were significant amounts of variation in measurement: the percent changes of the mean (SD) in the longest diameter was 5.8% (23.8) and 25.0% (117.9) in volume

This project seeks to extend that previous effort by altering the reading paradigm to more realistically reflect conditions encountered in clinical trials and clinical practice. In the proposed project, readers would be allowed to read cases side by side and would be allowed access to both visual results of previous markings as well as the quantitative results of those measurements (diameter, volume, etc.). In order to reduce the potential of bias from reading repeat CT scans in a side by side paradigm, we propose to introduce cases that do have some change in them; in this way, readers will not have the expectation that all cases will have "no change".

Primary goals and objectives

The primary goal of this project is to perform the analysis necessary to assess the minimum detectable change using reading conditions which are more realistic in clinical trials and clinical reading environments. **The minimum change will be analyzed for just one measurement method (volumetric measurements) and only Inter-reader variability will be assessed.**

Updated Study Design

Three major updates to the study design have been planned:

1. **Due to budgetary and time constraints, only volumetric measurements will be made (no single longest diameter, no bidimensional diameter)**
2. **We will use an equal number of "no change" and "change" cases**
3. **Because of changes in software, etc, each case will have two reads performed by each reader:**
    a. **An initial read of just the Timepoint 1 exam for all cases (even for the "no change" cases that were read in the Original 1B study)**
    b. **A subsequent read of each case in the "side by side" reading paradigm where readers are allowed to see both Timepoint 2 exam and the Timepoint 1 exam as well as their markups and measured values for the Timepoint 1 exam alone.**

Cases: we plan to use

1. **32 coffee break cases from the original 1B study (thin section, no change cases)**
2. **20 "change cases" that we acquired from the RIDER database**

These additional cases will be randomly introduced into the case mix so that readers will not have an expectation that all cases are "no change" cases. Some of these will be shown as "tumor progression" cases where the lesion will be larger in the second exam than in the first exam. Others will be shown as "responding to treatment" cases where the lesion will be smaller in the second exam than in the first exam (because we won't show any dates, we can simply do this by showing later exams as the first exam in the case of a true disease progression case).

This means we will have:
5 readers
52 cases (32 "no change" cases and 20 "change" cases)
2 time points of readings of each case by each reader
(520 reading events = 5 readers x 52 cases x 2 readings/case)

Analysis Plan

1. **Minimum detectable change under "side by side" review paradigm**
    a. **5 readers, 32 no change cases**
    b. **Compare volumes of lesion at timepoint 2 to volume of lesion at timepiont 1**
    c. **Inter-reader variability**
2. **Comparison to previous results of 1B with "independent" review paradigm**
3. **Analysis of inter-reader varability**
4. **Report the proportion and a 95% CI in the percent difference of "no-change", where the cases meet the condition of QIBA profile.**
5. **Descriptive statistics and comparison of "change" vs. "no-change" cases.**

Deliverables and Timetable [must include intermediate measureable milestones.]

1. Experimental design with side by side comparisons allowed, which will be presented to QIBA VolCT 1B group for approval. (2012 Q1; Feb 27, 2012)
2. Data analysis plan presented to QIBA VolCT 1B group for approval (2012 Q1; Approx. March 31)
3. Collection of "Change" cases (from RIDER database) (2012 Q1, Approx March 31
4. Actual data collection including image markup by readers completed. (2012 Q2; Approx June 30)
5. Completion of data analysis, including but not limited to:
    a. Investigation into minimum detectable change using revised reading paradigm (2012 Q3; Approx July 31, 2012)
    b. Inter-reader variability analysis (2012 Q3; Approx July 31, 2012)

6. Internal summary report of data analysis for QIBA members (2012 Q3; Approx Sept 30, 2012)
7. Submission of results to conferences (e.g. RSNA, SPIE) for presentation (2012 Q3; Approx Sept 30, 2012)
8. Submission of peer-reviewed publications based on results (2012 Q3; Approx Sept 30, 2012)