A Retrospective Exploratory Study of the Variability of Radiologists Measurements in a Selected Subgroup of Subjects Enrolled in a Clinical Trial

Robert W. Ford [1], Kathy Zhou Ph. D. [2], Robert R. Ford M.D.[1]

[1] RadPharm Imaging Core Lab (Princeton, NJ), [2] Department of Public Health Weill Medical College (New York, NY)

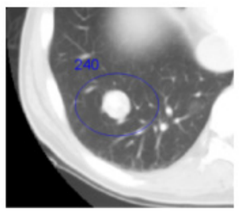
Background Information

- Independent centralized review of imaging studies is an established means of validating data used in support of corporate "go-no go" decisions, regulatory approval, and post marketing claims
- This study was performed to investigate and assign a concise value(s) to the inherent human variability in radiological measurements, despite best efforts to standardize assessment using Response Evaluation Criteria in Solid Tumors (RECIST)

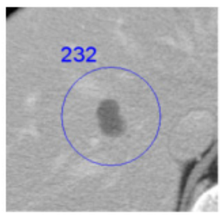
Materials and Methods

- An experienced radiologist (24 years) and a research intern reviewed digital Computerized Tomographic (CT) scans from 31 subjects to identify 150 primary and metastatic tumors. Tumors were categorized into 4 categories.

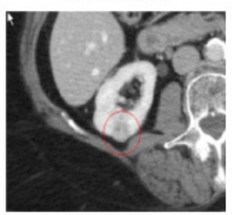
Tumor Category	Tumor Description
1	Defined Edge/Radially Symmetric
2	Defined Edge/Irregular Shape
3	Blurred Edge/Radially Symmetric
4	Blurred Edge/Irregular Shape



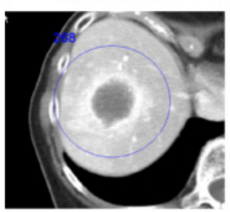
Category 1



Category 2



Category 3



Category 4

Training The Radiologist Readers

- Each radiologist reader attended a training session so that the study was performed consistently
- There was a presentation covering an example of lesion measurement and the correct method of recording obtained data into the form provided
- Also included in this presentation were the software operation and lesion measurement guidelines

Lesion Measurement

- At the time of the read, each radiologist was provided with written directions for software operation and Lesion Measurement
- Fifteen radiologists independently measured each tumor
- The zoom function was allowed to be used freely while window and level were held constant for each lesion
- The axial location of each lesion was specified to the radiologist by an image number
- The lesion to be measured was identified by a numbered annotation encircling the lesion that did not describe its geographic boundaries
- The lesions were presented to each reader in a categorically arbitrary sequence
- Each reader measured the longest dimension of each lesion in the axial plane using electronic calipers
- All scans were digital and thus required no pixel calibration (Calibration factors were encoded in the DICOM header)
- All readers used the same validated DICOM based software application

Statistical Analysis

- Statistical analysis was performed using mixed effects modeling to partition error according to its contributing factors
- Mixed Effects Modeling allows for the analysis of heterogeneous data

Results

Explanation of Terms

- *Total Error* – Accounts for all the variance seen in the measurements of the lesions
- *Between Reader Error* – The portion of the variance seen in the measurements that the radiologist readers are responsible for
- *Random Error* – The partition of the total error from all other contributing sources

Significant Data

- The between reader error accounted for 4.15% percent of total errors for the length
- The total error is smallest for category 1 and largest for category 4 as expected

	No of Tumors	Between Reader Error	Random Error	Total Error	Reader Error/Total Error
Overall	150	0.32 (0.14, 0.77)	7.51 (7.08, 8.01)	7.83	4.15%
Cat. 1	54	0.19 (0.08, 0.45)	1.25 (1.14, 1.39)	1.45	13.4%
Cat. 2	40	0.26 (0.08, 0.76)	4.67 (4.16, 5.24)	4.93	5.28%
Cat. 3	16	0.44 (0.14, 1.39)	3.92 (3.24, 4.75)	4.36	10%
Cat. 4	40	0.81 (0.25, 2.56)	19.98 (17.72, 22.47)	20.79	3.90%

Conclusions

- The [(low reader error) / (total error)] suggests that the radiologist readers were only responsible for a small portion of the total error found in this measuring process
- Due to the fact that the intra-reader variability is usually lower than the inter-reader variability one can speculate that these 15 radiologists reading one lesion would be comparable to 1 radiologist reading the same lesion 15 times

Continuing Research

- The structure and implementation of the project facilitates continuing research
- Analyze intra-reader variability by repeating the process at a later date
- Repeating the project will also help to strengthen conclusions made based on the current data
- Other functions in the software could be used as the variable
- This results database can now be used to qualify new radiologist readers and retest previously qualified readers