



# Evaluation of Number of Target Lesions to Analyze in Time to Progression by RECIST

Mithat Gönen<sup>1</sup>, Lawrence Schwartz<sup>2</sup>, Robert Ford<sup>3</sup>

Departments of Epidemiology and Biostatistics<sup>1</sup>, Radiology<sup>2</sup> Memorial Sloan-Kettering Cancer Center, New York, NY

<sup>3</sup>RadPharm, Princeton, NJ

## ABSTRACT

**Background:** RECIST criteria were designed to evaluate tumor shrinkage and response to therapy by measurement of multiple target lesions, evaluation of non target and new lesions. There is considerable controversy surrounding the optimal number of lesions to assess response, with RECIST requiring the measurement of up to 10 target lesions. These guidelines were set up to evaluate the endpoint of best overall response. Increasingly, time to progression has become an important endpoint in oncology trials. We evaluated the optimal number of lesions to measure to accurately and reproducibly assess time to progression.

**Methods:** We evaluated target lesions metastases in 1205 patients (**to be checked**) enrolled in three Phase III clinical trials. All patients underwent CT at baseline and standard follow up scans until progression. Target lesions were measured unidimensionally and response was assessed according to RECIST by 2 independent Radiologists. A total of 5190 (**to be checked**) target lesions were assessed. Response was calculated according to the rules of target lesions (up to 10) by RECIST, utilizing the 2 largest lesions and randomly selecting 2 target lesions.

**Results:** Using the 2 largest lesions, time to progression was concordant in 83% of cases. The 2 Radiologists determined the two same largest lesions in 89% of cases. Since the determination of the largest or the same target lesions is not always possible or performed, a random selection of 2 target lesions demonstrated a 76% concordance in the time to progression.

**Conclusions:** Measurement of time to progression may have a greater degree of variability than measurement of best overall response and therefore measurement of minimal selected lesions will lead to a great variability in response assessment.

## INTRODUCTION

RECIST criteria are commonly used to determine response to therapy, in multiple solid tumors. While RECIST criteria were established principally for the best overall response criteria, it is utilized for time to progression as well. In this study we demonstrate the level of discordance in determining time to progression using RECIST

## DATA DESCRIPTION

Three Phase III clinical trials are used for illustration

|                                       | Trial I | Trial II | Trial III | Total |
|---------------------------------------|---------|----------|-----------|-------|
| Number of Patients                    | 105     | 307      | 685       | 1097  |
| Number of Lesions                     | 890     | 2197     | 5944      | 9031  |
| Average Number of Lesions Per Patient | 8.5     | 7.2      | 8.7       | 8.2   |

Most Common Sites of Disease

| Site        | Trial I | Trial II | Trial III | Total |
|-------------|---------|----------|-----------|-------|
| Lung        | 35.6%   | 3.9%     | 29.1%     | 24.8% |
| Liver       | 12.6%   | 46.1%    | 9.2%      | 17.4% |
| Peritoneum  | 4.8%    | 28.7%    | 3.2%      | 8.7%  |
| Mediastinum | 9.6%    | 1.6%     | 12.3%     | 9.7%  |

Most Common Exam Types

| Exam            | Trial I | Trial II | Trial III | Total  |
|-----------------|---------|----------|-----------|--------|
| Chest CT        | 61.1%   | 5.8%     | 59.6%     | 48.7%  |
| Abdominal CT    | 32.7%   | 75.5%    | 36.5%     | 44.0%  |
| Pelvic CT       | 4.2%    | 16.3%    | 3.4%      | 8.7%   |
| Number of Exams | 10,644  | 15,205   | 47,459    | 73,308 |

## METHODS

Response is assessed in three different ways:

**All lesions:** This serves as the gold standard

- Largest lesions:** Largest two lesions only are considered. Repeated for largest three and four. Represents the "best case" scenario since largest lesions are determined after all lesions are measured and assumes multiple readers would select the same lesions
- Random lesions:** Randomly selected two lesions are considered. Repeated for randomly selected three and four. Represents the "worst case" scenario since selection of lesions to be included in the response assessment is usually more-informed than random, but may represent the true clinical scenario when lesions are of similar size and Radiologist work independently

## LARGEST LESIONS

### Percent Discordant

| Number of Lesions Included | Trial I | Trial II | Trial III |
|----------------------------|---------|----------|-----------|
| 2                          | 9.5%    | 2.0%     | 0.8%      |
| 3                          | 5.7%    | 0.9%     | 0.7%      |
| 4                          | 4.8%    | 0.5%     | 0.2%      |

## LARGEST LESIONS

### Discrepancy with Gold Standard (months)

| Number of Lesions Included | Trial I | Trial II | Trial III |
|----------------------------|---------|----------|-----------|
| 2                          | 8.2     | 2.11     | 2.96      |
| 3                          | 14.54   | 2.111    | 3.08      |
| 4                          | 5.69    | 3.91     | 1.78      |

## RANDOM LESIONS

### Percent Discordant

| Number of Lesions Included | Trial I | Trial II | Trial III |
|----------------------------|---------|----------|-----------|
| 2                          | 13.3%   | 5.3%     | 0.6%      |
| 3                          | 12.3%   | 3.9%     | 0.6%      |
| 4                          | 9.5%    | 1.3%     | 0.6%      |

## RANDOM LESIONS

### Discrepancy with Gold Standard (months)

| Number of Lesions Included | Trial I | Trial II | Trial III |
|----------------------------|---------|----------|-----------|
| 2                          | 2.91    | 1.41     | 2.43      |
| 3                          | 5.53    | 1.38     | 2.43      |
| 4                          | 7.91    | 2.07     | 2.43      |

## CONCLUSIONS

- Even in best case scenario two largest lesions (RECIST) rule may be discordant as often as 10%
- This percent of discordant cases decreases as greater number of lesions are measured however there is a more variability on effect on the discordance in months
- As the number of lesions measured increase, reproducibility of determination of progression improves
- Determining time to progression using RECIST has greater variability than determining response.