

## Presentation 2020: ASA Modeling and Simulating Reader Studies

[Back to Presentations and Publications Main Page](#)

Final presentation:

- [ASA\\_ModelingandSims.pdf](#) (3 MB, uploaded by Katherine N Elfer 3 years 5 months ago)

Related pre-print submitted to Statistics in Biopharmaceutical Research:

- <https://arxiv.org/abs/2107.08891>

## American Statistical Association

The ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop

### Modeling and Simulating Reader Studies to Support the Evaluation of Image-Based Algorithms

- 1:30 PM, Wednesday, 23 September 2020
- Virtual

#### Authors

- Brandon D. Gallas (1)
- Si Wen (1)

(1) FDA/CDRH/OSEL/DIDSR, Silver Spring, MD, US

#### Abstract

A growing part of the medical device portfolio of CDRH includes image-based detection (e.g., find the tumor) and classification algorithms (e.g., classify an abnormality as benign or malignant). Whatever the health condition, imaging technology, or algorithm architecture (neural networks, random forests, regressions), submissions of the “software as a medical device” often include a reader study, a study in which clinicians make evaluations with and without the algorithm. Comparing the evaluations against a reference truth, we can compare the performance impact of the algorithm. The statistical analysis of such studies is not trivial since it is well known that there is a range of skill among clinicians and their evaluations are noisy. Furthermore, the studies often have multiple clinicians evaluating the same cases, leading to correlations in the data. FDA guidance recommends an MRM (multi-reader multi-case) analysis paradigm in which a reader-averaged performance metric is analyzed (variance estimates, confidence intervals, and p-values) to account for the variability (and correlations)

from readers and cases. To support such analyses, we have developed, published, and shared on [GitHub](#) statistical methods and software, data, and examples. Such development relies on simulations of MRMC studies to validate the statistical methods. In this talk, we will discuss reader studies, performance metrics, and the corresponding MRMC structures of uncertainty. We will present a simulation model that has served us well in validating MRMC analyses of detection and classification metrics. To address studies of algorithms that yield quantitative values and the within- and between-clinician agreement of such values, we have been developing new MRMC methods that analyze differences in quantitative values. To support this work, we are investigating and will present a new simulation model that better represents such data.