Statistical Analysis Plan

July 16, 2018

FMSU2017-003: A Feature Comparison Study to Evaluate the Modified Processing of Fujifilm's ASPIRE Cristalle with Digital Breast Tomosynthesis Option as Compared to the Original Processing

FUJIFILM Medical Systems U.S.A., Inc.

NCT03586583

July 16, 2018

Statistical Considerations for
FUJIFILM Medical Systems, U.S.A., Inc. protocol FMSU2017-003

1. Statistical Analysis Plan:

The primary objective of the study is to test the non-inferiority of the modified processing in each of 7 general mammographic features. The proportion of cases where the readers judged the modified image processing as non-inferior will be calculated. Non-inferiority is defined as a score of equivalence between modified and original processing or a score indicating the modified image processing is superior. For each reader, the proportion of cases judged as non-inferior with modified processing will be calculated for each feature and each view. A pooled estimate will also be calculated where the results over all readers and views will be combined. 95% percentile bootstrap CIs will be constructed for the pooled proportions of cases judged as non-inferior with the modified processing. These CIs will be reported for each feature.

The null hypothesis is that the pooled proportion of cases judged as non-inferior with modified image processing is ≤ 0.50, versus the alternative hypothesis that the pooled proportion is >0.50. If the 95% CI did not contain values ≤ 0.50, then the null hypothesis will be rejected. These hypotheses will be evaluated for each of the seven general mammographic features. If the null hypothesis is rejected for each of the seven general features, then it will be concluded that the readers' subjective evaluation of the features on the modified imaging is non-inferior to the original imaging.

2. Sample Size Justification:

The following assumptions were made for determining sample size for the primary objectives:

- One breast per subject (two views)
- Each of six readers scores an independent sample of cases; thus, there is no correlation between readers' interpretations.
- 50% of cases have dense breasts
- $\geq$80% power is required
- 2.5% type I error rate (Note that hypothesis rejection depends on the lower bound of the 95% CI being greater than 0.50.)

Based on these assumptions, sample size was determined as follows:

$$M = \frac{\{1.96 \times \sqrt{0.5(1-0.5)} + 0.84\sqrt{p(1-p)}\}^2}{[p-0.5]^2}$$

where p is the proportion of comparisons judged as non-inferior under the alternative hypothesis. The sample size is given by N=M/(# readers). Table 1 summarizes the number of subjects needed per reader as a function of the proportion of cases judged as non-inferior with the modified processing.

Table 1. Number of Subjects Needed Per Reader for Tests of Non-inferiority

# readers	Proportion of Comparisons Judged as Non-inferior with Modified Processing (DVIIm plus ISR)			
	0.55	**0.56**	**0.57**	**0.60**
6	131	91	67	33

A study with 600 total subjects (i.e. 100 cases evaluated per reader) was proposed. This sample size would provide slightly more than 80% power to reject the primary null hypothesis, as long as ≥56% of cases are judged non-inferior.

3. Randomization

Cases will be randomly selected from Fujifilm's existing library (collected under protocol FMSU2013-004A). A stratified randomization will be used where the strata and proposed sampling proportions are summarized in Table 2. Within each stratum, cases will be randomly selected so that approximately 10% would have extremely dense breasts, 40% would have heterogeneously dense breasts, 40% would have scattered fibroglandular breasts, and 10% would have fatty breasts.

Table 2. Sampling Plan

Stratum, defined by Reference Standard*	# available subjects	Overall sample composition	# cases per reader	Total # cases
Cancer: (N=229)				
Architectural Distortion	21	38%	3-4	21
Microcalcification	86		14-15	86
Asymmetry	19		3-4	19
Mass	95		15-16	95
Other	8		1-2	8
Biopsy Benign: (N=457)				
Architectural Distortion	19	32%	3-4	19
Microcalcification	212		10	60
Asymmetry	43		7-8	43
Mass	158		9	54
Other	25		2-3	15

Recalled/No Biopsy	98	10%	10	60
Negative	234	10%	10	60
Unknown	209	10%	10	60
Total	**1227**	**100%**	**100**	**600**

Because there are subjects with multiple lesions, cases in the table were organized first by the presence/absence of a malignant lesion, and then by presence/absence of an AD (and possibly another lesion), microcalcification, asymmetry, mass, and other, respectively.

For each reader, a randomization list with 100 unique cases will be created consisting of four blocks:

- Block 1: 50 breasts from 50 subjects where the reader scored the CC view,
- Block 2: 50 breasts from a second sample of 50 subjects where the reader scored the MLO view,
- Block 3: same 50 breasts as in Block 1, but re-ordered and scored with the MLO view, and
- Block 4: same 50 breasts as in Block 2, but re-ordered and scored with the CC view.