## Appendix E1

## Site Qualification Process

Each MRI system used in the ACRIN 6698 study was required to pass a DW MRI qualification process incorporating assessment of both phantom and patient scans, described in detail in ACRIN 6698 study materials (18). Briefly, sites were provided a pair of 1.5 L cylindrical ice-water filled breast DW MRI phantoms each containing a central 29 mm diameter measurement tube prefilled with distilled water. Site investigators were instructed to fill the space surrounding the measurement tube with an ice-water slurry and allow one hour for thermal equilibrium before scanning (35). Phantom pairs were scanned using the site's bilateral breast coil according to a specified protocol. DW MRI phantom images were analyzed at the site qualification laboratory (by T.L.C., over 20 years experience in quantitative DW MRI analysis) for the following performance metrics (with target values) within fixed $1 \times 2$ cm$^2$ rectangular regions-of-interest (ROIs) drawn in the central measurement tubes: ADC bias error ($< 10\%$); ADC random error ($< 5\%$); ADC b-value dependence ($< 2\%$); ADC left-vs-right difference ($< 5\%$); and signal-to-noise ratio at b = 800 sec/mm$^2$ ($> 75:1$). In addition, key acquisition parameters (eg, b-value, repetition time, echo time, acquisition voxel size, etc) were determined from phantom image headers to assess protocol conformance. Failure to meet a single metric target did not automatically fail the system, rather system "pass/fail" was judged by the quantity and severity failures, overall phantom image quality and conformance to study protocol. If a given site/system failed DW MRI phantom tests, it was provided corrective recommendations and invited to resubmit phantom images. After passing DW MRI phantom tests, each site was requested to submit two representative clinical breast MRI examinations for image quality review by the ACRIN 6698 protocol team to assess DW MRI features related to overall parenchymal signal strength and uniformity, level of artifact, and fat-suppression quality. Prior to site study initiation, online training sessions were held with site PIs and research staff to educate them to acceptable data quality in terms of SNR, adequate fat suppression, and minimal artifacts and distortions. Examples were shown for both good quality cases along with common breast DW MRI image quality issues and methods to address these.

## Reference

35. Malyarenko D, Galbán CJ, Londy FJ, et al. Multi-system repeatability and reproducibility of apparent diffusion coefficient measurement using an ice-water phantom. J Magn Reson Imaging 2013;37(5):1238–1246 .

**Table E1: Standardized ACRIN 6698 MRI Acquisition Parameters**

| Parameter | T2-weighted | Diffusion-weighted | T1-weighted |
|---|---|---|---|
| Sequence type | FSE or STIR | DW SE-EPI | GE |
| 2D or 3D sequence | 2D | 2D | 3D |
| Slice orientation | Axial or sagittal | Axial | Axial |
| Laterality | Bilateral | Bilateral | Bilateral |
| Frequency direction | A/P | A/P | A/P |
| Phase direction | R/L (axial) | R/L | R/L |

| | S/I (sagittal) | | |
|---|---|---|---|
| *FOV–frequency | 260–360 mm (axial) 180–220 mm (sagittal) | 260–360 mm | 260–360 mm |
| *FOV-phase | 260–360 mm (axial) 180–220 mm (sagittal) | 300–360 mm | 260–360 mm |
| Matrix–frequency (acquired) | 256–512 | 128–192 | 384–512 |
| Matrix–phase (acquired) | $\geq 256$ | 128–192 | $\geq 256$ |
| Reconstruction Matrix | 512 x 512 | 256 x 256 | 512 x 512 |
| In-plane resolution | $\leq 1.4$ mm | 1.7–2.8 mm | $\leq 1.4$ mm |
| Fat-suppression | Active fat-sat recommended | Active fat-sat | Active fat-sat recommended |
| Repetition Time (TR) | 2000–10000 ms | $\geq 4,000$ ms | 4–10 ms |
| Echo Time (TE) | 70–140 msec (STIR 70 ms) | Minimum | Minimum |
| Echo Train Length | $\leq 16$ | N/A | N/A |
| Inversion Time (TI; STIR sequence) | 170 msec (1.5T) 230 msec (3.0T) | N/A | N/A |
| Flip Angle | 90 degrees | 90 degrees | 10–20 degrees |
| Readout Bandwidth (per pixel) | N/A | N/A | Maximum |
| b values | N/A | 0, 100, 600, 800 s/mm$^2$ | |
| Slice thickness (acquired) | $\leq 4$ mm | 4–5 mm | $\leq 2.5$ mm |
| Number of slices | Variable; complete bilateral coverage | Variable; complete bilateral coverage | $\geq 60$; complete bilateral coverage |
| Slice Gap | $\leq 1.0$ mm | No gap | No gap |
| Parallel imaging factor | $\leq 2$ | $\geq 2$ | $\leq 2$ |
| No. of excitations or averages | $\leq 2$ | $\geq 2$ | $\leq 2$ |
| k-space ordering | N/A | N/A | -k to +k standard, noncentric |
| Sequence acquisition time | $\leq 7$ minutes | $\leq 5$ minutes | 80 sec $\leq$ scan time $\leq$ 100 sec |
| Total postcontrast imaging duration | N/A | N/A | $\geq 8$ minutes following injection |

Abbreviations: FSE = fast spin echo, STIR = short tau inversion recovery, DW = diffusion-weighted, SE = spin echo, EPI = echo planar imaging, GE = gradient echo, 2D = two dimensional, 3D = three-dimensional, A/P = anterior-posterior, R/L = right-left, S/I = superior-inferior, FOV = field-of-view, N/A = not applicable.

## Table E2: MRI Systems Utilized for the Eligible and Analysis Cohorts

| | Eligible ($n = 388$) | | Analysis Set ($n = 242$) | |
|---|---|---|---|---|
| VENDOR, n (%) | | | | |
| GE | 275 | (71) | 174 | (72) |
| Siemens | 59 | (15) | 36 | (15) |
| Philips | 47 | (12) | 32 | (13) |
| Missing | 7 | (1.8) | 0 | (0.0) |
| FIELD STRENGTH, n (%) | | | | |
| 1.5 tesla | 253 | (65) | 169 | (70) |
| 3.0 tesla | 128 | (33) | 73 | (30) |
| Missing | 7 | (1.8) | 0 | (0.0) |

## Table E3: Performance of Tumor ΔFTV for Predicting pCR Stratified by Cancer Subtype

| | ΔFTV (%) | | | | | | |
|---|---|---|---|---|---|---|---|
| | pCR | | NonpCR | | AUC | 95% CI | P value |
| | n | Mean ± SD | n | Mean ± SD | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| HR-/HER2- | | | | | | | |
| Early-treatment/3 weeks | 29 | −39 ± 46 | 48 | −21 ± 53 | 0.61 | (0.48, 0.74) | 0.09 |
| Midtreatment/12 weeks | 27 | −76 ± 34 | 44 | −68 ± 50 | 0.56 | (0.42, 0.70) | 0.41 |
| Posttreatment | 29 | −95 ± 8.8 | 42 | −86 ± 19 | 0.74 | (0.62, 0.86) | <0.001 |
| HR+/HER2- | | | | | | | |
| Early-treatment/3 weeks | 15 | −57 ± 28 | 83 | −27 ± 55 | 0.69 | (0.45, 0.78) | 0.006 |
| Midtreatment/12 weeks | 15 | −79 ± 54 | 77 | −70 ± 43 | 0.70 | (0.62, 0.89) | 0.013 |
| Posttreatment | 14 | −96 ± 6.1 | 78 | −82 ± 31 | 0.70 | (0.55, 0.87) | 0.007 |
| HR-/HER2+ | | | | | | | |
| Early-treatment/3 weeks | 18 | −76 ± 15 | 6 | −56 ± 49 | 0.54 | (0.16, 0.91) | 0.85 |
| Midtreatment/12 weeks | 18 | −93 ± 7.6 | 4 | −64 ± 36 | 0.71 | (0.27, 1.00) | 0.36 |
| Posttreatment | 17 | −94 ± 9.6 | 5 | −88 ± 13 | 0.69 | (0.40, 0.99) | 0.20 |
| HR+/HER2+ | | | | | | | |
| Early-treatment/3 weeks | 18 | −68 ± 20 | 24 | −53 ± 42 | 0.58 | (0.40, 0.76) | 0.37 |
| Midtreatment/12 weeks | 17 | −89 ± 14 | 24 | −87 ± 16 | 0.54 | (0.36, 0.73) | 0.64 |
| Posttreatment | 18 | −95 ± 6.4 | 22 | −88 ± 25 | 0.53 | (0.35, 0.72) | 0.74 |

Abbreviations: ΔFTV = change in functional tumor volume from pretreatment, pCR = pathologic complete response, AUC = area under the receiver operating characteristic curve, SD = standard deviation.