10.1148/radiol.2016161286

Appendix E1: Supplemental Methods and Results

Algorithm Training Procedure

The open-source LIBRA algorithm was trained to allow breast density estimation from sDMs by following the same procedure outlined by Keller et al in the original LIBRA publication (24). Briefly, a set of randomly selected 146 sDMs (n = 73 women, bilateral MLO-view images) with corresponding standard-dose images available was used as the training set. Breast PD estimates made by the LIBRA algorithm on each of the corresponding standard-dose images were used as the ground-truth PD estimates for training.

The breast region within each synthetic mammogram was identified by using LIBRA's automated segmentation, and adaptive fuzzy c-means clustering was applied to group breast tissue regions into areas of similar gray-level image intensity. The ground-truth PD estimates obtained from the corresponding standard-dose digital mammograms were then utilized to label pixels within the N-highest–intensity clusters within the segmented breast region as "dense" and the remainder as "non-dense," where N is determined by the algorithm on a per-image basis so as to minimize the error between the obtained estimates and the ground-truth PD for that image.

Subsequently, within each cluster, a set of 86 features, as previously described by Keller et al (24), were computed, including features related to the image acquisition physics, global histogram statistics and texture features, cluster-specific histogram statistics and texture features, and between-cluster differences in features. Linear stepwise feature selection was then used to identify those features most significantly associated with PD for inclusion into a support vector machine (SVM) classifier. Finally, the SVM classifier with the selected features was trained on the entire data set of 146 images, on the basis of the cluster ground-truth assignment described above.

Validation of Training

To validate the trained version of the algorithm for the sDMs and verify its performance, radiologists' manual estimates of PD were also obtained from the synthetic mammograms by a fellowship-trained breast imaging radiologist (E.F.C., with more than 20 years of clinical experience) using the Cumulus semiautomated software. Pearson correlations between the synthetic mammogram estimates and radiologist estimates were calculated to assess the trained synthetic mammogram method's accuracy, and the paired Student t test was used to detect the presence of any systematic difference in the density scores between the automated synthetic density estimates and the radiologists' estimates made on the same synthetic images.

Results

The automated breast PD estimates made by LIBRA on the synthetic mammograms were strongly correlated with density estimates made on standard-dose mammograms (r = 0.95, P < .001) in the training set. In addition, the automated density estimates from the synthetic mammograms were also strongly correlated with radiologist-provided ground-truth PD scores from the synthetic images (r = 0.89, P < .001). The paired t test showed that automated PD

estimates from synthetic mammograms were 1.5% higher, on average, than estimates made on the standard-dose mammograms ($P = .01$) and 5.7% higher, on average, than the radiologists' estimates made on the synthetic mammograms ($P < .001$), indicating good agreement as compared with broadly acceptable inter- and intrareader agreement (20,35,37).