This is the supplemental document for the MICCAI manuscript titled "Preoperative prediction of lymph node metastasis from clinical DCE MRI of the primary breast tumor using a 4D CNN". In the main paper the architecture of the best performing model, a novel hybrid 4D CNN model that integrates 4D DCE MRI and clinical characteristics of the patient is described. In Table 2 the paper compares the performance of this model to 3 models that use: (1) clinical information alone, (2) 2D images (axial orientation), and (3) 3D image volumes that have been similarly image-preprocessed as described in the paper. In this supplemental a description of the models of the three additional experiments is provided below.

To predict metastasis from clinical information alone, an **XGboost** model is trained which employs the random forest gradient boosting approach as described in detail in [1]. This model has two main hyperparameters: the number of trees, *n_tree,* and the maximum tree depth, *max_depth*. These govern the statistical complexity of the model. A grid search is employed to optimize these hyperparameters over the ranges: $10 \leq n\_tree \leq 60$ and $3 \leq max\_depth \leq 5$. The model with the best validation AUC is selected and used to obtain the result on the held out test set, a set which is not used for training nor model selection.

To predict metastasis from 2D input images a 2D CNN is used. *First, ROIs delineating the primary tumor were draw by board-certified radiologists approximately 3 min post contrast. Then, 2D square 120x120 patches were cropped around the centroid of the ROI.*I. Fig. S1 visualizes the top 2D model from the validation data and how the data dimensions change after each layer. Model hyperparameters are tuned through empirical testing. The model has nine 2D convolution layers (red pyramids). The 2nd, 3rd, and 4th are followed by a 2x2 maxpooling layer. The result from the final convolutional layer is flattened and a fully connected layer with 36 units is used to learn the mapping to the two-category output. All layers use batchnormalization.

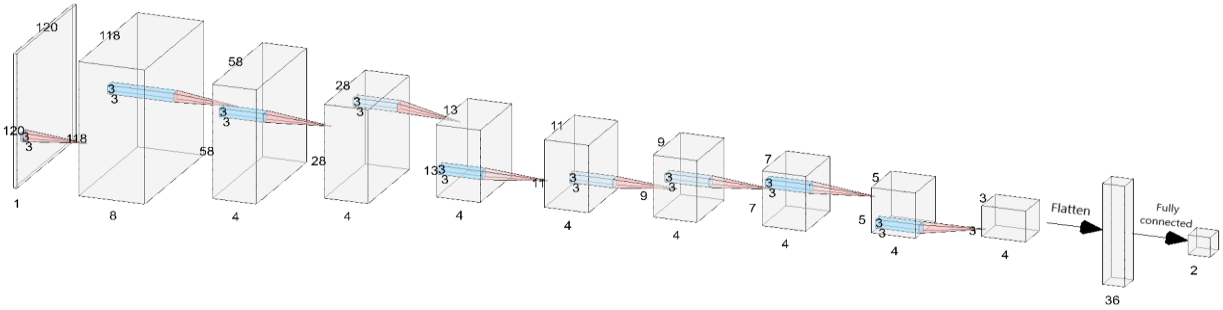

Fig. S1. The 2D CNN model architecture. The model consists of 9 convolutional layers (red pyramids) followed by 1 fully connected layer (horizontal arrow, on right). One-hot encoded output layer predicts probability of nodal metastasis and no nodal metastasis.

To predict metastasis from 3D input volumes a 3D CNN is trained**.** The architecture of the top model is shown in Fig. S2. CNN Model hyperparameters are similarly tuned through empirical testing. This model has 4 convolutional layers followed by maxpooling. The result from the convolutional layers is flattened and used as input to a 3 layer dense feedforward neural network that learns the two-category output.

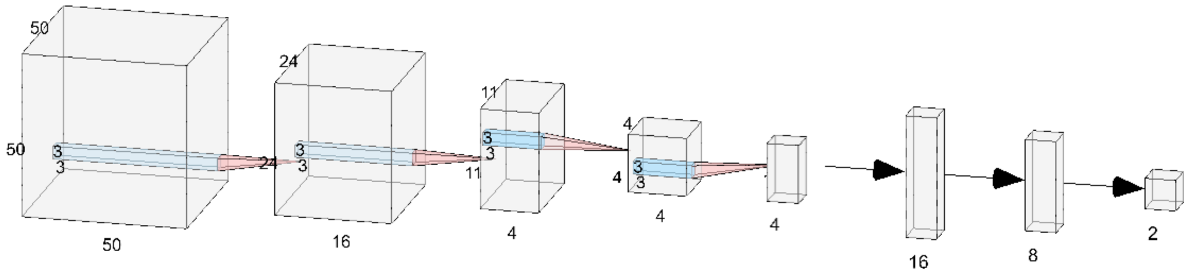

Fig. S2. The 3D CNN model architecture. The top model consists of 4 convolutional layers (red pyramids) followed by 3 fully connected layers (horizontal arrows). One-hot encoded output layer predicts probability of nodal metastasis and no nodal metastasis.

Supplemental References:

[1] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794).