**Supplementary Table 1.** Layer by layer description of the tissue segmentation network architecture, with layers grouped into a total of 10 blocks. The down sampling portion (Blocks 1-5), inspired by VGG-Net, has 10 convolutional layers and 3 maximum pooling layers. Our up sampling portion (Blocks 6-8), inspired by U-Net, consists of 3 blocks that each has a bilinear resize operation followed by a convolutional layer. Block 9 is a 1 x 1 convolutional layer, which serves as the convolutional equivalent of a fully connected layer. Block 10 is our output softmax layer.

|  |  |
| --- | --- |
| **Block Number** | **Layer Specifications** |
| **1** | 128 x 128 x 64 convolutional layers (kernel = 3 x 3)128 x 128 x 64 convolutional layers (kernel = 3 x 3)2 x 2 maximum pooling |
| **2** | 64 x 64 x 128 convolutional layers (kernel = 3 x 3)64 x 64 x 128 convolutional layers (kernel = 3 x 3)2 x 2 maximum pooling |
| **3** | 32 x 32 x 256 convolutional layers (kernel = 3 x 3)32 x 32 x 256 convolutional layers (kernel = 3 x 3)2 x 2 maximum pooling |
| **4** | 16 x 16 x 512 convolutional layers (kernel = 3 x 3)16 x 16 x 512 convolutional layers (kernel = 3 x 3) |
| **5** | 16 x 16 x 512 convolutional layers (kernel = 3 x 3)16 x 16 x 512 convolutional layers (kernel = 3 x 3) |
| **6** | 32 x 32 x 512 bilinear resize32 x 32 x 256 convolutional layers (kernel = 3 x 3) |
| **7** | 64 x 64 x 256 bilinear resize32 x 32 x 192 convolutional layers (kernel = 3 x 3) |
| **8** | 128 x 128 x 192 bilinear resize128 x 128 x 64 convolutional layers (kernel = 3 x 3) |
| **9** | 128 x 128 x 4 convolutional layers (kernel = 1 x 1) |
| **10** | 128 x 128 x 4 softmax |

**Supplementary Table 2.** The amount of epithelium (%) and breast cancer risk stratified by benign breast disease (BBD) histological subtypes, parity, menopausal status, and BMI. Data were analyzed using unconditional logistic regression models adjusting for the matching factors to estimate odd ratios (ORs) and 95% confidence intervals (CI).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | **Quartile 1** | **Quartile 2** | **Quartile 3** | **Quartile 4** | ***p-*trend** |
| **BBD histological subtype** |  |  |  |  |  |
| Non-proliferative | Cases/Controls, *n* | 13/83 | 12/82 | 12/83 | 26/83 |  |
|  | Model | Ref | 1.00 (0.42,2.34) | 0.91 (0.38,2.17) | 1.86 (0.84,4.26) | 0.07 |
| Proliferative without atypia | Cases/Controls, *n* | 23/161 | 32/161 | 32/161 | 61/162 |  |
| Model | Ref | 1.28 (0.71,2.33) | 1.14 (0.63,2.09) | 1.92 (1.11,3.40) | **0.01** |
| Atypical hyperplasia | Cases/Controls, *n* | 18/39 | 24/39 | 22/39 | 18/39 |  |
| Model | Ref | 1.07 (0.48,2.39) | 1.08 (0.48,2.40) | 0.70 (0.29,1.67) | 0.38 |
| **Parity** |  |  |  |  |  |
| Nulliparous women | Cases/Controls, *n* | 2/27 | 5/26 | 8/27 | 13/27 |  |
| Model | Ref | 2.64 (0.50,20.15) | 4.76 (1.00,35.31) | 7.36 (1.58,54.68) | **0.01** |
| Parous women | Cases/Controls, *n* | 54/255 | 53/255 | 68/255 | 90/255 |  |
| Model | Ref | 0.95 (0.63,1.46) | 1.13 (0.75,1.70) | 1.35 (0.90,2.03) | 0.08 |
| **Menopausal Status** |  |  |  |  |  |
| Pre-menopausal women | Cases/Controls, *n* | 34/170 | 43/169 | 62/170 | 58/170 |  |
| Model | Ref | 1.24 (0.75,2.06) | 1.70 (1.06,2.77) | 1.46 (0.89,2.40) | 0.17 |
| Post-menopausal women | Cases/Controls, *n* | 11/91 | 18/91 | 16/91 | 23/92 |  |
| Model | Ref | 1.68 (0.76,3.87) | 1.53 (0.68,3.59) | 2.05 (0.95,4.62) | 0.11 |
| **Body Mass Index** |  |  |  |  |  |
| <25 | Cases/Controls, *n* | 34/160 | 33/160 | 54/160 | 65/161 |  |
| Model | Ref | 0.89 (0.52,1.53) | 1.32 (0.81,2.20) | 1.39 (0.84,2.33) | 0.09 |
| 25-30 | Cases/Controls, *n* | 11/76 | 17/75 | 21/76 | 26/76 |  |
| Model | Ref | 1.56 (0.69,3.67) | 1.94 (0.87,4.52) | 2.25 (1.02,5.25) | 0.07 |
| >30 | Cases/Controls, *n* | 11/43 | 7/43 | 4/43 | 10/44 |  |
| Model | Ref | 0.54 (0.18,1.54) | 0.36 (0.09,1.16) | 0.85 (0.31,2.32) | 0.94 |

The amount of epithelial tissue (%) was categorized into quartiles as defined by the distribution among the controls. The median value for each quartile was included as a continuous variable in the unconditional logistic regression together with matching factors to obtain the *p-*trend value (Wald test).

**Supplementary Table 3.** Tissue composition (%) and breast cancer risk defined by tumor estrogen receptor (ER) expression. Data were analyzed using polytomous logistic regression models to estimate odd ratios (ORs) and 95% confidence intervals (CI).

|  |  |  |
| --- | --- | --- |
|  | **ER-positive breast cancer (*n*=185)** | **ER-negative breast cancer (*n*=51)** |
|  | **T1** | **T2** | **T3** | ***p-trend*** | **T1** | **T2** | **T3** | ***p-trend*** |
|  |  |  |  |  |  |  |  |  |
| **Epithelium (%)** |  |  |  |  |  |  |  |  |
| Cases/Controls, *n* | 50/374 | 56/384 | 79/374 |  | 10/374 | 16/384 | 25/374 |  |
| Model 1 | Ref | 1.06 (0.70,1.60) | 1.45 (0.97,2.17) | 0.05 | Ref | 1.48 (0.66,3.34) | 2.05 (0.93,4.5) | 0.07 |
| Model 2 | Ref | 0.88 (0.57,1.34) | 1.24 (0.82,1.89) | 0.18 | Ref | 1.46 (0.64,3.32) | 2.05 (0.91,4.58) | 0.08 |
| Model 3 | Ref | 1.01 (0.64,1.59) | 1.32 (0.84,2.07) | 0.16 | Ref | 1.21 (0.52,2.85) | 1.88 (0.83,4.26) | 0.10 |
| **Fibrous stroma (%)** |  |  |  |  |  |  |  |  |
| Cases/Controls, *n* | 55/374 | 55/384 | 75/374 |  | 16/374 | 16/384 | 19/374 |  |
| Model 1 | Ref | 0.94 (0.63,1.41) | 1.30 (0.88,1.94) | 0.20 | Ref | 0.84 (0.41,1.72) | 0.94 (0.46,1.90) | 0.85 |
| Model 2 | Ref | 0.85 (0.56,1.29) | 1.24 (0.83,1.86) | 0.31 | Ref | 0.82 (0.40,1.68) | 0.92 (0.45,1.89) | 0.82 |
| Model 3 | Ref | 0.86 (0.56,1.34) | 1.18 (0.76,1.82) | 0.45 | Ref | 0.85 (0.40,1.78) | 0.85 (0.40,1.78) | 0.66 |
| **Fat (%)** |  |  |  |  |  |  |  |  |
| Cases/Controls, *n* | 83/374 | 52/384 | 50/374 |  | 22/374 | 18/384 | 11/374 |  |
| Model 1 | Ref | 0.62 (0.42,0.92) | 0.62 (0.41,0.95) | **0.04** | Ref | 0.94 (0.48,1.82) | 0.70 (0.31,1.57) | 0.39 |
| Model 2 | Ref | 0.61 (0.41,0.91) | 0.67 (0.43,1.04) | 0.10 | Ref | 0.93 (0.48,1.81) | 0.70 (0.31,1.59) | 0.40 |
| Model 3 | Ref | 0.63 (0.41,0.95) | 0.70 (0.44,1.12) | 0.17 | Ref | 1.05 (0.53,2.07) | 0.75 (0.32,1.78) | 0.53 |

Each tissue region was categorized into tertiles (T1, T2, and T3) as defined by the distribution among the controls. Model 1 adjusted for matching factors. Model 2 adjusted for matching factors and BBD histological subtypes. Model 3 adjusted for matching factors, BBD histological subtypes, parity, menopausal status, and BMI. The median value for each quartile was included as a continuous variable in the polytomous logistic regression for Model 1, 2, or 3, to obtain the p-trend value (Wald test).

**Supplementary Table 4.** The association of tissue composition (%) and breast cancer risk, with each tissue region analyzed as a continuous variable per 10% change and two of the three tissue regions simultaneously included in the model. Unconditional logistic regression models were used to estimate odd ratios (ORs) and 95% confidence intervals (CI).

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Epithelial tissue** | **Stroma tissue** | **Fat tissue** |
| **Per 10% increase in epithelium, or fibrous stroma substituting for fat**  |
| Model 1 | 1.31 (1.08,1.59) | 1.10 (0.97,1.25) | - |
| Model 2 | 1.28 (1.05,1.57) | 1.07 (0.94, 1.22) | - |
| Model 3 | 1.30 (1.05, 1.61) | 1.03 (0.90, 1.19) | - |
| **Per 10% increase in epithelium, or fat substituting for fibrous stroma** |
| Model 1 | 1.19 (0.99, 1.43) | - | 0.91 (0.80, 1.03) |
| Model 2 | 1.20 (0.99, 1.45) | - | 0.94 (0.82, 1.07) |
| Model 3 | 1.26 (1.03, 1.54) | - | 0.97 (0.84, 1.12) |
| **Per 10% increase in fibrous stroma, or fat substituting for epithelium** |
| Model 1 | - | 0.84 (0.70, 1.02) | 0.76 (0.63, 0.93) |
| Model 2 | - | 0.83 (0.69, 1.01) | 0.78 (0.64, 0.96) |
| Model 3 | - | 0.79 (0.65, 0.97) | 0.77 (0.62, 0.95) |

Model 1 adjusted for matching factors. Model 2 adjusted for matching factors and BBD histological subtypes. Model 3 adjusted for matching factors, BBD histological subtypes, parity, menopausal status, and BMI.