**Additional file 1: Table S1. Pack-years of smoking**

|  |
| --- |
| To calculate smoking pack-years |
| Divide the number of cigarettes smoked per day by 20 (the number of cigarettes in a pack).Then multiply by the number of years smoked. |
| example 1 | (60 cigarettes/day ÷ 20 cigarettes/pack) × 10 years = 30 pack-years |
| example 2 | (40 cigarettes/day ÷ 20 cigarettes/pack) × 20 years = 40 pack-years |
| example 3 | (10 cigarettes/day ÷ 20 cigarettes/pack) × 40 years = 20 pack-years |

**Table 2.** **Clinicopathological features of 149 patients who were treated with preoperative chemotherapy.**

|  |  |
| --- | --- |
| Parameters (*n* = 149) | Number of patients (%) |
| Age (years old) | 56 (24–75) |
| Tumour size (mm) | 27.6 (10.2 – 98.0) |
| Skin infiltrationNegative / Positive | 134 (89.9 %) / 15 (10.1 %) |
| Lymph node metastasis　N0 / N1 / N2 / N3 | 51 (34.2 %) / 54 (36.2 %) / 29 (19.5 %) / 15 (10.1 %) |
| HER2 Negative / Positive | 87 (58.4 %) / 62 (41.6 %) |
| Ki67Negative / Positive | 23 (15.4 %) / 126 (84.6 %) |
| ORR Non-Responders / Responders | 11 (7.4 %) / 138 (92.6 %) |
| pCR Negative / Positive | 75 (50.3 %) / 74 (49.7 %) |
| Recurrence Negative / Positive | 123 (82.6 %) / 26 (17.4 %) |
| TILs Low / High | 58 (38.9 %) / 91 (61.1 %) |
| Smoker No / Yes | 105 (70.5 %) / 44 (29.5 %) |
| Pack-years of smokers | 20 (2.5 – 135) |
| Pack-years Low / High | 106 (71.1 %) / 43 (28.9 %) |

HER: human epidermal growth factor receptor. ORR: objective response rate. pCR: pathological complete response. TILs: tumour-infiltrating lymphocytes.

**Table 3. Comparison of** **clinicopathological features by subtype.**

|  |  |  |  |
| --- | --- | --- | --- |
| Parameters |  | Intrinsic subtype | *p* value |
| HER2-enriched breast cancer(*n* = 62) | Triple-negative breast cancer(*n* = 87) |
| Age (years old) ≤ 56 > 56 | 26 (41.9 %)36 (58.1 %) | 49 (56.3 %)38 (43.7 %) | 0.085 |
| Tumour size (mm) ≤ 27.6 > 27.6 | 30 (48.4 %)32 (51.6 %) | 45 (51.7 %)42 (48.3 %) | 0.690 |
| Skin infiltration Negative Positive | 54 (87.1 %)8 (12.9 %) | 80 (92.0 %)7 (8.0 %) | 0.335 |
| Lymph node status Negative Positive | 25 (40.3 %)37 (59.7 %) | 26 (29.9 %)61 (70.1 %) | 0.188 |
| Ki67 Negative  Positive | 14 (22.6 %)48 (77.4 %) | 9 (10.3 %)78 (89.7 %) | 0.042 |
| ORR Non-Responders Responders | 1 (1.6 %)61 (98.4 %) | 10 (11.5 %)77 (88.5 %) | 0.023 |
| pCR Negative Positive | 26 (41.9 %)36 (58.1 %) | 49 (56.3 %)38 (43.7 %) | 0.085 |
| Recurrence Negative Positive | 55 (88.7 %)7 (11.3 %) | 68 (78.2 %)19 (21.8 %) | 0.096 |
| TILs Low High | 20 (32.3 %)42 (67.7 %) | 38 (43.7 %)49 (56.3 %) | 0.206 |
| Pack-years Low High | 40 (64.5 %)22 (35.5 %) | 66 (75.9 %)21 (24.1 %) | 0.134 |

HER: human epidermal growth factor receptor. ORR: objective response rate. pCR: pathological complete response. TILs: tumour-infiltrating lymphocytes.

**Table 4.** **Difference in clinicopathological features due to pack-years.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Parameters |  | Smoker | *p* value | Pack-years | *p* value |
| Yes(*n* = 44) | No(*n* = 105) | High(*n* = 43) | Low(*n* = 106) |
| Age (years old) ≤ 56 > 56 | 24 (54.5 %)20 (45.5 %) | 51 (48.6 %)54 (51.4 %) | 0.509 | 24 (55.8 %)19 (44.2 %) | 51 (48.1 %)55 (51.9 %) | 0.398 |
| Tumour size (mm) ≤ 27.6 > 27.6 | 25 (56.8 %)19 (43.2 %) | 50 (47.6 %)55 (52.4 %) | 0.309 | 25 (58.1 %)18 (41.9 %) | 50 (47.2 %)56 (52.8 %) | 0.228 |
| Skin infiltration Negative Positive | 40 (90.9 %)4 (9.1 %) | 94 (89.5 %)11 (10.5 %) | 0.799 | 40 (93.0 %)3 (7.0 %) | 94 (88.7 %)12 (11.3 %) | 0.428 |
| Lymph node status Negative Positive | 18 (40.9 %)26 (59.1 %) | 33 (31.4 %)72 (68.6 %) | 0.269 | 18 (41.9 %)25 (58.1 %) | 33 (31.1 %)73 (68.9 %) | 0.214 |
| Ki67 Negative  Positive | 6 (13.6 %)38 (86.4 %) | 17 (16.2 %)88 (83.8 %) | 0.696 | 6 (14.0 %)37 (86.0 %) | 17 (16.0 %)89 (84.0 %) | 0.752 |
| Intrinsic subtype HER2-enriched Triple-negative | 22 (50.0 %)22 (50.0 %) | 40 (38.1 %)65 (61.9 %) | 0.181 | 22 (51.2 %)21 (48.8 %) | 40 (37.7 %)66 (62.3 %) | 0.134 |
| ORR Non-Responders Responders | 3 (6.8 %)41 (93.2 %) | 8 (7.6 %)97 (92.4 %) | 0.866 | 3 (7.0 %)40 (93.0 %) | 8 (7.5 %)98 (92.5 %) | 0.905 |
| pCR Negative Positive | 17 (38.6 %)27 (61.4 %) | 58 (55.2 %)47 (44.8 %) | 0.065 | 16 (37.2 %)27 (62.8 %) |  59 (55.7 %)47 (44.3 %) | 0.042 |
| Recurrence Negative Positive | 40 (90.9 %)4 (9.1 %) |  83 (79.0 %)22 (21.0 %) | 0.083 | 39 (90.7 %)4 (9.3 %) |  84 (79.2 %)22 (20.8 %) | 0.096 |
| TILs Low High | 13 (29.6 %)31 (70.5 %) | 45 (42.9 %)60 (57.1 %) | 0.075 | 12 (27.9 %)31 (72.1 %) | 46 (43.4 %)60 (56.6 %) | 0.043 |

HER: human epidermal growth factor receptor. ORR: objective response rate. pCR: pathological complete response. TILs: tumour-infiltrating lymphocytes.

**Table 5.** **Univariate and multivariate analysis with respect to disease-free survival and overall survival.**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Univarite analysis |  | Multivariate analysis |
| Parameters | Hazard ratio | 95% CI | *p* value |  | Hazard ratio | 95% CI | *p* value |
| disease-free survival |
| Age at opetation (yr) ≤ 56 > 56 | 0.614 | 0.269-1.335 | 0.220 |  |  |  |  |
| Tumour size (mm) ≤ 27.6 > 27.6 | 1.137 | 0.525-2.503 | 0.744 |  |  |  |  |
| Skin infiltration Negative Positive | 1.556 | 0.455-4.067 | 0.440 |  |  |  |  |
| Lymph node status Negative Positive | 2.440 | 0.933-8.343 | 0.071 |  | 1.677 | 0.617-5.859 | 0.331 |
| Ki67 Negative  Positive | 0.394 | 0.180-0.926 | 0.034 |  | 0.770 | 0.321-1.942 | 0.568 |
| Intrinsic subtype HER2-enriched Triple-negative | 1.884 | 0.828-4.823 | 0.135 |  |  |  |  |
| ORR Non-Responders Responders | 0.083 | 0.035-0.210 | <0.001 |  | 0.154 | 0.059-0.426 | 0.001 |
| Pathological response Non-pCR pCR | 0.203 | 0.068-0.499 | <0.001 |  | 0.381 | 0.118-1.059 | 0.065 |
| TILs Low High | 0.252 | 0.107-0.553 | 0.001 |  | 0.424 | 0.167-1.032 | 0.059 |
| Pack-years Low High | 0.434 | 0.127-1.134 | 0.092 |  | 0.567 | 0.160-1.555 | 0.289 |
| overall survival |
| Age at opetation (yr) ≤ 56 > 56 | 0.508 | 0.175-1.338 | 0.172 |  |  |  |  |
| Tumour size (mm) ≤ 27.6 > 27.6 | 1.123 | 0.429-2.993 | 0.811 |  |  |  |  |
| Skin infiltration Negative Positive | 1.939 | 0.446-5.977 | 0.335 |  |  |  |  |
| Lymph node status Negative Positive  | 2.778 | 0.781-17.657 | 0.125 |  |  |  |  |
| Ki67 Negative  Positive | 0.638 | 0.234-2.023 | 0.419 |  |  |  |  |
| Intrinsic subtype HER2-enriched Triple-negative | 1.610 | 0.597-5.060 | 0.357 |  |  |  |  |
| ORR Non-Responders Responders | 0.077 | 0.026-0.238 | <0.001 |  | 0.451 | 0.044-0.520 | 0.004 |
| Pathological response Non-pCR pCR | 0.183 | 0.042-0.561 | 0.002 |  | 0.282 | 0.062-0.953 | 0.041 |
| TILs Low High | 0.357 | 0.129-0.929 | 0.035 |  | 0.634 | 0.212-1.861 | 0.403 |
| Pack-years Low High | 0.554 | 0.128-1.700 | 0.325 |  |  |  |  |

OS: overall survival. CI: confidence intervals. HER: human epidermal growth factor receptor. ORR: objective response rate. pCR: pathological complete response. TILs: tumour-infiltrating lymphocytes.