**Supplementary Table 2 The baseline characteristics of colon cancer patients in the training and validation sets.**

|  |  |  |
| --- | --- | --- |
| Clinical characteristics | TCGA cohort (n=452) | GEO cohort\* (n=787) |
| Age (median, range) | 69 (31-90) | 68 (22-97) |
| Gender (%)  Female  Male | 214 (47.3%)  238 (52.7%) | 361 (45.9%)  426 (54.1%) |
| T (%)  Tis  T1  T2  T3  T4  Unknow | 1 (0.2%)  10 (2.2%)  77 (17.0%)  308 (68.1%)  56 (12.4%)  0 (0%) | 0 (0%)  12 (1.5%)  47 (6.0%)  376 (47.8%)  119(15.1%)  233 (29.6%) |
| N (%)  N0  N1  N2  N3  Unknow | 269 (59.5%)  103 (22.8%)  80 (17.7%)  0 (0%)  0 (0%) | 306 (38.9%)  136 (17.3%)  100 (12.7%)  6 (0.8%)  239 (30.4%) |
| M (%)  M0  M1  MX  Unknow | 334 (73.9%)  62 (13.8%)  49 (10.8%)  7 (1.5%) | 491 (62.4%)  61 (7.8%)  2 (0.3%)  233 (29.6%) |
| Stage (%)  Stage I  Stage II  Stage III  Stage IV  Unknow | 76 (16.8%)  178 (39.3%)  125 (27.7%)  62 (13.7%)  11 (2.4%) | 63 (8.0%)  334 (42.4%)  279 (35.5%)  111 (14.1%)  0 (0%) |
| Status (%)  Alive  Dead | 364 (80.5%)  88 (19.5%) | 507 (64.4%)  280 (35.6%) |

\*A GEO cohort integrating GSE17536, GSE17537, GSE17538 and GSE39582