의학박사 학위논문

에스트로겐 수용체 양성, 표피 성장 인자 2 수용체 음성 여성 유방암 환자에서 종양의 ^{18}F-fluorodeoxyglucose 대사와 암생존과의 연관성

Association between tumor ^{18}F-fluorodeoxyglucose metabolism and survival in women with estrogen receptor-positive, ERBB2-negative breast cancer

울산대학교 대학원

의 학 과

박설훈

Association between tumor ^{18}F-fluorodeoxyglucose metabolism and survival in women with estrogen receptor-positive, ERBB2-negative breast cancer

지 도 교 수 문대혁

이 논문을 의학박사 학위 논문으로 제출함

2021 년 8 월

울 산 대 학 교 대 학 원

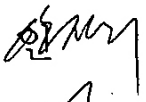
의 학 과

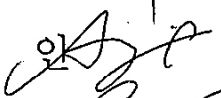
박설훈

박설훈의 의학박사학위 논문을 인준함

심사위원 류진숙 

심사위원 문대혁 

심사위원 안진희 

심사위원 나세정 

심사위원 이효상 

울 산 대 학 교 대 학 원

2021 년 8 월

ABSTRACT

Purpose: We examined whether tumor ^{18}F-fluorodeoxyglucose (^{18}F-FDG) metabolism is associated with distant relapse-free survival (DRFS) in women with estrogen receptor (ER)-positive, ERBB2-negative breast cancer.

Materials and methods: This is a multi-cohort study using public and clinical cohorts of breast cancer patients. Public data with microarray expression up to May 2019 were collected from three referral centers in the USA and Japan. A cohort from Asan Medical Center, Korea (Asan cohort), recruited between November 2007 and December 2014, was also included. Women with ER-positive, ERBB2-negative breast cancer who received anthracycline-based neoadjuvant chemotherapy were included in this study. The primary outcome of this study was DRFS. Twelve multigene scores (including the ^{18}F-FDG signature score) for the public cohorts and maximum standardized uptake value (SUV) of ^{18}F-FDG positron emission tomography (PET) for the Asan cohort were measured. Additionally, a separate public cohort (lung cancer patients with ^{18}F-FDG PET and gene expression profiles)

was analyzed for correlation between tumor ^{18}F-FDG metabolism and ^{18}F-FDG signature score.

Results: The analysis included 394 women (297 with survival data) from the public cohorts and 466 from the Asan cohort. The median (interquartile range) age was 49 (43–58) and 45 (39–51) years, respectively. The median (interquartile range) follow-up period for patients without distant metastasis or death was 3.2 (2.3–4.4) for the public cohorts and 6.2 (5.3–7.6) years for the Asan cohort, respectively. Both ^{18}F-FDG signature and maximum SUV were associated with clinical characteristics. In a separate public cohort of lung cancer patients, ^{18}F-FDG signature score showed a moderate correlation with maximum SUV. The ^{18}F-FDG signature score of public cohorts was independently associated with DRFS (HR, 4.85; 95% confidence interval [CI], 1.57–11.05; $P = 0.006$), and the prognostic discrimination (C-index: 0.70; 95% CI, 0.62–0.78) was comparable with those of other multigene scores. In the Asan cohort, multivariable analysis of tumor ^{18}F- FDG metabolism showed that the middle and high tertiles of maximum SUV were prognostic for DRFS (Ter1 vs. Ter2 = HR, 2.26; 95%

CI, 1.17–4.39; $P = 0.02$: Ter1 vs. Ter 3 = HR, 2.93; 95% CI, 1.55–5.54; $P = 0.001$). The 8–year DRFS was 90.7% (95% CI, 85.5–96.1%) for the low tertile and 73.7% (95% CI, 68.0–79.8%) for the middle and high tertile values.

Conclusion: ^{18}F-FDG PET may assess the risk of distant metastasis and death in ER-positive, ERBB2-negative patients.

Keywords: Breast cancer, estrogen receptor, glucose metabolism, survival, positron emission tomography

Contents

List of figures

List of tables

Introduction

Estrogen receptor (ER)-positive, ERBB2-negative breast cancer accounts for 60–70% of breast cancers[1]. Although patients are treated with curative intent, >20% of patients experience relapse within 10–15 years and die from metastatic disease[2]. The recommended systemic treatment for patients with operable but high-risk tumors, and those with locally advanced disease includes chemotherapy, followed by endocrine therapy. However, the diverse molecular characteristics of patients lead to diversity in the response to systemic therapies[3,4]. ER-positive, ERBB2-negative breast cancer does not benefit from systemic chemotherapy to the same extent as other subtypes[5], and not all ER-positive breast cancers respond optimally to endocrine therapy. Identifying patients with newly diagnosed cancer who are at a high risk of relapse despite standard chemo- and endocrine therapy could help select patients who require more effective treatments or provide a basis for recommending participation in clinical trials[6].

The diversity in transcriptional programs accounts for much of the biological heterogeneity of breast cancer[7]. The luminal epithelial-specific genes, including ER and

proliferation genes, are the main gene clusters that are differentially expressed among intrinsic subtypes[3]. Likewise, multigene prognostic assays primarily rely on ER and proliferation-related gene expression[8-16]. A meta-analysis of gene expression profiles from large cohorts revealed that the capacity of prognostic signatures depends mainly on the detection of proliferation activity, and that ER expression status may contain only indirect information about prognosis[16]. The Warburg effect of aerobic glycolysis, a key metabolic hallmark of cancer, fuels cell growth and proliferation. Positron emission tomography (PET) using ^{18}F-fluorodeoxyglucose (^{18}F-FDG) allows visualization of the increased glucose metabolism in malignant tumors. In ER-positive, ERBB2-negative breast cancers, the maximum standardized uptake value (SUV) of ^{18}F-FDG is associated with poor prognostic factors, including progesterone receptor status[17], histologic grade, Ki-67 expression[18], and 21-gene recurrence score[17]. ^{18}F-FDG PET provides prognostic information in a non-invasive manner and with a low overall cost when previously performed for staging. However, there is limited information on the ability of ^{18}F-FDG PET to predict distant relapse-free survival (DRFS) and overall survival (OS).

The main objective of this study was to examine the association of tumor ^{18}F-FDG metabolism with DRFS in multiple cohorts of ER-positive, ERBB2-negative breast cancer patients treated with anthracycline-based neoadjuvant chemotherapy (NCT) followed by adjuvant endocrine therapy. A pooled analysis of publicly available prognostic data sets with gene expression profiles was performed for computation of ^{18}F-FDG and multigene prognostic signature scores in the same patients[19]. A cohort from our institution was also analyzed. Administration of chemotherapy before surgery has an advantage over adjuvant chemotherapy in that it allows the differentiation of predictive markers of the response to chemotherapy from prognostic factors. Therefore, we also examined whether the ^{18}F-FDG signature score, maximum SUV, and multigene prognostic signature scores were associated with pathological complete response (pCR) to NCT.

Materials and methods

Study design

This is a multi-cohort study composed of three cohort groups; 1) publicly available cohorts of breast cancer for calculation of multigene signature scores, 2) a clinical cohort of Asan Medical Center for evaluation of ^{18}F-FDG tumor metabolism, and 3) a separate public cohort of lung cancer for evaluating correlation between FDG values. The study design is summarized in Fig. 1.

Study setting and patients of publicly available cohorts

Publicly available cohorts with microarray gene expression profiles and patient outcome information were selected from the Gene Expression Omnibus and ArrayExpress up to May 2019 using the keywords "breast neoplasm," and "neoadjuvant chemotherapy". Seven hand-searched articles were analyzed to identify publicly available gene expression profiling data and clinical outcomes of patients with ER-positive breast cancer[20-26]. Data series profiled on GeneChip® Human Genome U133A and U133 Plus 2.0 arrays (Affymetrix Inc., Santa Clara,

USA) were included to minimize the potential bias introduced by merging data series with different platforms[27]. After reviewing all retrieved summaries and citations, data series analyzing participants who received anthracycline-based chemotherapy were included. Data series that had no information on pathological complete response or distant relapse-free survival, or those that included only ER-negative participants, were excluded. An additional inclusion criterion at the participant level was estrogen receptor-positive, ERBB2-negative disease assessed by immunohistochemistry.

Study setting and patients of Asan cohort

A clinical cohort of women with ER-positive, ERBB2-negative breast cancer who received anthracycline-based NCT between November 2007 and December 2014 at Asan Medical Center (Asan cohort), located in Seoul, Republic of Korea, was also included. Follow-up ended on December 31, 2019. Risk factors were examined in relation to outcomes that had already occurred at the start of the study. The institutional review board of Asan Medical Center approved the study protocol and waived the requirement for informed consent. The

study was performed following the Declaration of Helsinki and institutional guidelines. Enrolled patients had American Joint Committee on Cancer stage II or III tumors (tumor stage ≥T2) with a histological type of invasive ductal carcinoma. All patients underwent ^{18}F-FDG PET/CT and received at least one cycle of NCT. Patients were excluded if they had a prior history of cancer or bilateral breast cancer. The number of patients enrolled during the study period determined the sample size of the Asan cohort.

Gene expression analysis of multigene signature scores

Data processing and quantification were performed using R statistical software version 3.6.0 (R Foundation for Statistical Computing, Vienna, Austria, available at https://www.r-project.org). The single-channel array normalization algorithm implemented in the SCAN.UPC package (version 2.26.0) from Bioconductor (http://bioconductor.org/biocLite.R) was used to normalize the data of each series downloaded from the Gene Expression Omnibus repository (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi)[28]. Probes were mapped to Entrez Gene unique integers using the Brainarray chip definition files (version

22) available at http://brainarray.mbni.med.umich.edu/Brainarray/Database/CustomCDF/genOmic_curated_CDF.asp[29]. The four normalized datasets were merged using the sva package (version 3.32.1) to remove the potential batch effect across the data sets[30]. An empirical Bayes algorithm was used in the merging process[31].

The ^{18}F-FDG signature incorporated 75 genes from the glycolysis, pentose phosphate, and carbon fixation metabolic pathways[19], and each gene was weighted based on the degree of differential expression between the ^{18}F-FDG high and ^{18}F-FDG low primary breast tumor samples. Metabolic pathway-based ^{18}F-FDG signature scores were obtained using weighted gene voting[32].

Eleven multigene signatures with a reported association with prognosis were analyzed, including the 21-gene recurrence[8], 70-gene signature[9], 50-gene signature[21], 12-gene risk[33], 76-gene signature[34], genomic grade index[12], CIN70 signature[13], PTEN signature[14], proliferation gene module[15], ESR1 signature[15], and single sample predictor of intrinsic molecular subtype[21]. The signature scores were calculated, and the intrinsic molecular subtypes (luminal A, luminal B, HER2-enriched, basal-like, and normal-like based on the

prediction analysis of microarray 50 assay) were determined for each sample using the Bioconductor genfu package (version 2.16.0)[35]. For CIN70 and PTEN signature scores that were not included in the genfu package, a previously published method was used[13,14]. The signature genes of 21-gene recurrence, 76-gene signature, genomic grade index, proliferation gene module, and ESR1 were mapped to probe set ID and to Entrez Gene unique integers for the rest of the signatures. Only genes that could be mapped to probe set or Entrez Gene IDs were used. Each signature score was rescaled by designating 2.5 and 97.5 percentile values as 0 and 1, respectively.

^{18}F-FDG PET/CT of the Asan cohort

^{18}F-FDG PET/CT imaging was performed from the skull base to the upper thigh at 50–70 min after intravenous administration of 5.2–7.4 MBq/kg (0.14–0.2 mCi/kg) of ^{18}F-FDG[36]. The median blood glucose level before ^{18}F-FDG injection was 102 mg/dl (IQR, 94–109). Two board-certified nuclear medicine physicians who were blinded to patient outcomes drew a volume of interest on the primary breast cancer or metastatic lymph nodes and assessed the

maximum SUV of ^{18}F-FDG uptake. The maximum standardized uptake values were harmonized across various PET/CT scanners (Biograph Sensation 16 and Biograph TruePoint 40, Siemens Healthineers, Knoxville, TN, USA; Discovery STE 8, Discovery PET/CT 690, and Discovery PET/CT 710, GE Healthcare, Milwaukee, WI, USA) without partial volume correction[37]. The recovery coefficient profiles of variable-sized hot cylinders of the American College of Radiology-approved PET phantoms (i.e., Esser phantom) were matched[38,39]. Annual ^{18}F water cylinder phantom-based cross-calibration between PET and dose calibrator assured the uniform standardized uptake value of 1.0 between PET scanners[39]. Maximum SUV was analyzed as a continuous measurement or as a categorical estimate by grouping the patients according to the median or tertile value.

Lung cancer microarray data with ^{18}F-FDG PET (A separate public cohort)

Publicly available breast cancer cases with both ^{18}F-FDG PET and gene expression profile data could not be identified. Alternatively, 27 matched ^{18}F-FDG PET images and gene expression profiles of lung cancer patients were retrieved from The Cancer Imaging Archive

(http://www.cancerimagingarchive.net) and The Cancer Genome Atlas (deposited at the Genomic Data Commons Data Portal; https://portal.gdc.cancer.gov/exploration), respectively. The inclusion criteria for cases were samples containing ≥50% tumor cells and a metabolic tumor volume of ≥4.2 mL[40,41]. Raw mRNA counts were retrieved from the Genomic Data Commons Data Portal, and normalized and log-transformed using the weighted trimmed mean of M-values algorithm implemented in the edgeR Bioconductor package (version 3.26.8)[42]. The ^{18}F-FDG signature scores were calculated as described in the previous section. The normalized maximum standardized uptake value (SUV) of the primary tumor was calculated as the maximum SUV divided by the mean SUV of the normal liver tissue within a spherical volume of interest with a diameter of 3 cm using PET-VCAR software on an Advantage Workstation (version 4.3; GE Healthcare). For correlation analysis, the ^{18}F-FDG signature scores were rescaled to the same mean and standard deviation of the normalized maximum SUV.

Variables

The primary outcome measure of this study was DRFS. The secondary outcome was OS. DRFS was defined as the interval from the date of initial diagnostic biopsy (the public cohorts) or NCT (the Asan cohort) to the diagnosis of distant metastasis or death from breast cancer, non-breast cancer, or unknown causes[25,43]. OS was measured until the date of death from any cause. pCR was defined as the absence of residual invasive cancer on hematoxylin and eosin-stained samples of the complete resected breast specimen and all sampled regional lymph nodes[24,25,44,45].

Factors considered potential predictors of DRFS and OS that were prespecified in the study protocol were age, tumor stage, clinical lymph node stage, histologic grade, ER Allred score, progesterone receptor status, and Ki-67 expression, as previously described[24,25,44-47]. All variables were retrieved from the public databases and the electronic medical records of Asan Medical Center. Predictors that had continuous values or belonged to three or more categories were dichotomized based on commonly used cut-off values that are relevant for prognosis as follows[48-50]: age (20–50 vs. >50), tumor stage (T1–2 vs. T3–4), clinical N stage (N0 vs. N1–3), histologic grade (1 or 2 vs. 3), ER score (3–6 vs. 7–8), and Ki-67 expression

(<20% vs. ≥20%).

Statistical analysis

Data were expressed as the median and IQR for continuous variables or numbers (%) for categorical variables. A two-sided *P* value of < 0.05 was considered significant. Categorical variables were compared using the χ^2 test. For continuous variables, the Mann-Whitney U or Kruskal-Wallis test with post hoc Dunn's test was used. Associations were analyzed using linear regression or Spearman rank correlation. Survival curves were estimated using the Kaplan-Meier method and compared with the log-rank test.

Logistic regression analysis was used to determine the predictors of pCR, and Cox proportional hazard regression analysis was applied to identify independent predictors of DRFS and OS based on a predetermined statistical plan. Harrell's concordance index (C-index) was used to compare the discriminatory capacity of the multigene signature models. Optimism-corrected C-index was obtained by bootstrapping validation (100 bootstrap resamples). Patients were censored to the date of the last disease assessment. For the

multivariable analysis, a model was constructed including clinical and pathological predictors that were associated with the outcome at $P \leq 0.10$. Post hoc analysis was performed to assess whether the overall association was consistent across all subgroups studied. For logistic and Cox regression, linearity, potential influential observations, and multicollinearity were assessed. The proportional hazard assumption was tested with Schoenfeld's residual test and Schoenfeld's partial residual plots. A restricted cubic spline plot was used for detecting nonlinearity and deviance residual for examining influential observations. Whether an overall association was consistent across all subgroups of patients categorized according to variables for survival analysis was examined. Estimates were presented with 95% CI.

Missing clinical and pathological data of the public and Asan cohorts were imputed using the multiple imputation method (m=10). Ten imputed data sets were generated, and estimates from the multiple imputed data sets were combined according to Rubin's rule[51]. The robustness of the results was demonstrated by performing a complete case analysis as a sensitivity analysis of the imputation model.

Statistical tests were performed using R statistical software, version 3.6.1 (The R Foundation for Statistical Computing) and IBM SPSS Statistics for Windows, version 21 (IBM Corporation).

Results

Study cohorts and patients

The search identified 18 data series that met the inclusion criteria. Four data series with 383 participants were selected for analysis of pCR, and two series with 297 participants were selected for survival analysis[24,25,44]. The figure shows the flow of patients through the study (Fig. 2). The public cohorts included data from the University of Texas M.D. Anderson Cancer Center (USA), the Brown University Oncology Group (USA), and Osaka University Hospital (Japan).

The Asan cohort comprised 466 women with breast cancer (Fig. 3) with demographics typical of the neoadjuvant setting, and it included patients with regional lymph node disease or T3–4 tumors (Table 1). The median time between ^{18}F-FDG PET/computed tomography (CT) and the first date of NCT was 8 days (interquartile range [IQR], 6–10). All patients except five who were lost to follow-up (n = 4) or had ER-negative cancer after surgery (n = 1) received adjuvant tamoxifen or aromatase inhibitor therapy.

Multigene prognostic signature scores of the public patient cohorts

The ^{18}F-FDG signature score and other multigene scores from gene expression profiles were obtained for all patients from public cohorts. All genes included in the 21-gene recurrence, 12-gene risk, 76-gene signature, genomic grade index, chromosomal instability 70 genes (CIN70) signature, and proliferation gene module were mapped. However, we mapped 73 of 75 genes for the ^{18}F-FDG signature, 43 of 70 genes for the 70-gene signature, 43 of 50 genes for the 50-gene signature and a single sample predictor of the intrinsic molecular subtype, 151 of 205 genes for the phosphatase and tensin homolog (PTEN) signature, and 468 of 469 genes for the estrogen receptor 1 signature. There was no significant difference in ^{18}F-FDG signature scores among four publicly available breast cancer cohorts ($P = 0.89$, Fig. 4). Other multigene prognostic signatures did not differ significantly among the four groups ($P > 0.40$). The distribution suggests that the ^{18}F-FDG signature was a continuous variable (Fig. 5). Regarding intrinsic subtypes, 136 (34%) were luminal A, 152 (38%) were luminal B, 42 (11%) were *ERBB2*-enriched, 38 (10%) were basal-like, and 26 (7%) were normal-like breast cancer.

The ^{18}F-FDG signature score had a moderate to strong positive association with most multigene prognostic signature scores; it had a weak negative association with estrogen receptor 1 (ESR1) ($P < 0.001$, Fig. 6). The ^{18}F-FDG signature scores differed significantly according to the intrinsic molecular subtype. The luminal A and normal-like subtypes had a low ^{18}F-FDG signature score, whereas the luminal B, *ERBB2*-enriched, and basal-like subtypes had significantly higher ^{18}F-FDG signature scores than the luminal A subtype ($P < 0.001$, Fig. 7a). On the other hand, the luminal A and B subtypes had significantly higher ESR1 scores than other subtypes ($P < 0.001$, Fig. 7b). All other multigene prognostic signature scores showed the same distribution as the ^{18}F-FDG signature score according to the intrinsic subtype ($P < 0.001$, Fig. 8). There was a moderate association between the ^{18}F-FDG signature score and SUV in patients with lung cancer (Fig. 9).

Patient characteristics of the study cohorts

The median age of patients in the public cohorts was 49 years (IQR, 43–58). There was significant heterogeneity across the public cohorts regarding tumor stage ($P = 0.001$),

histologic grade ($P < 0.001$), and pCR ($P = 0.003$). The ^{18}F-FDG signature score was associated with clinical characteristics (Table 2).

The median age of the Asan cohort was 45 years (IQR, 39–51). The baseline values of the Asan cohort were significantly different from those of the public cohorts with respect to age ($P < 0.001$), tumor stage ($P = 0.03$), histologic grade ($P < 0.001$), progesterone receptor status ($P = 0.003$), and NCT response ($P < 0.001$). The maximum SUV ranged from 1.36 to 25.06, with a median value of 5.14 (Fig. 10), and was associated with clinical characteristics (Table 2). The clinical and pathological characteristics of patients who did not undergo ^{18}F-FDG PET/CT before NCT were not significantly different from those of patients who underwent ^{18}F-FDG PET/CT (Table 3).

Response to NCT

Of 383 patients in the public cohorts, 45 (12%) achieved a pCR. Among clinicopathological characteristics, histologic grade was associated with pCR, with a C-index of 0.69 (Table 4). All multigene prognostic signatures, including the ^{18}F-FDG score, were associated with pCR

(Table 4). Multivariable models combining each multigene score and histologic grade showed that most multigene prognostic signatures, including ESR1, were independently associated with pCR (Table 5), whereas the ^{18}F-FDG signature score was not. Of 460 patients in the Asan cohort who underwent surgery, 22 (5%) achieved a pCR (Table 6). Multivariable analysis including histologic grade and progesterone receptor showed that positivity for progesterone receptor was independently associated with pCR (odds ratio, 0.21; 95% confidence interval [CI], 0.08–0.51; $P = 0.001$).

Prognostic value

The median follow-up period for patients in the public cohorts without distant metastasis or death was 3.2 (IQR, 2.3–4.4) years. Distant metastasis or death did not occur in any of the 30 (0%) patients who achieved a pCR. Patients with pCR had a significantly longer DRFS ($P = 0.02$, Fig. 11). Among the clinical and pathological factors, clinical lymph node stage was significantly associated with DRFS (Table 7). ^{18}F-FDG and multigene prognostic signatures, except the 76-gene, CIN70, and proliferation gene module, were significantly associated

with DRFS (Table 7). Kaplan-Meier survival curves of DRFS according to tertiles of multigene signatures are shown in Fig. 12. Multivariable analysis combining each multigene prognostic signature with clinical lymph node stage showed that high ^{18}F-FDG, 21-gene, 70-gene, 50-gene, 12-gene, genomic grade index scores, and low ESR1 signature scores were independently associated with reduced DRFS (Table 8). Prognostic information from ^{18}F-FDG, ESR1, or their combination was comparable to that from other multigene prognostic signatures ($P > 0.10$, Table 9). After deleting cases with missing clinical N stage and histologic grade values, the results were consistent with the main prognostic data presented.

The median follow-up of patients in the Asan cohort without distant metastasis or death was 6.2 (IQR, 5.3–7.6) years. Distant metastasis or death did not occur in the 22 patients of the Asan cohort who achieved a pCR. Patients with pCR had significantly longer DRFS ($P = 0.04$, Fig. 13). Age, tumor stage, clinical N stage, Ki-67 expression, and maximum SUV were associated with DRFS (Table 10). Multivariable analysis showed that maximum SUV was independently associated with DRFS and OS when analyzed as a categorical estimate by grouping the patients according to the median or tertile value (Table 11). Ki-67 expression

was not independently associated with DRFS and OS regardless of whether we analyzed it as a continuous variable or several ordered categories (data not shown). The apparent C-index of the final model for DRFS was 0.72 (95% CI, 0.67–0.77). The optimism-corrected C-index was 0.70 (95% CI, 0.65–0.75). Kaplan-Meier estimates for patients who were free of distant metastasis or death were significantly different according to the maximum SUV ($P < 0.001$, Fig. 14). Post-hoc analysis of DRFS and OS showed that the overall association between maximum SUV and survival was consistent across most clinical and pathological subgroups large enough to provide sufficient outcome data (Fig. 15–16). When patients were stratified into subgroups according to independent prognostic factors, the increase in absolute risk of distant metastasis and death associated with high maximum SUV applied predominantly to patients with clinically node-positive disease (Fig. 17–18).

Discussion

In the present study, we showed that high tumor ^{18}F-FDG metabolism was associated with reduced DRFS and OS after adjusting for standard prognostic factors in patients with ER-positive, ERBB2-negative breast cancer treated with anthracycline-based NCT, followed by adjuvant endocrine therapy. The ^{18}F-FDG signature provided prognostic information on DRFS that was comparable to that obtained from validated multigene prognostic signatures. Tumor ^{18}F-FDG metabolism was significantly associated with standard prognostic factors and multigene prognostic signatures. This suggests that ^{18}F-FDG PET can help in early identification of a high-risk population in ER-positive, ERBB2-negative breast cancer patients and in establishing an appropriate treatment plan. To the best of our knowledge, this study is the first to demonstrate the value of tumor ^{18}F-FDG metabolism for the long-term prognosis of DRFS and OS in patients with ER-positive, ERBB2-negative breast cancer.

The results obtained from the public cohorts suggest that high ^{18}F-FDG and multigene prognostic signature scores are associated with a higher rate of pCR. However, the association of these factors with pCR is the opposite of their association with poor survival.

The number of patients who achieved a pCR was small, and the rate of poor prognosis was mainly determined by those who did not achieve pCR. High ^{18}F-FDG and multigene signature scores were associated with worse DRFS despite the higher chemotherapy sensitivity in these patients. Previous studies investigating ER-positive breast cancer obtained the same results for multigene prognostic signatures[21,25,52,53]. These observations highlight the fact that survival is influenced not only by chemotherapy response, but also by baseline biologic features and sensitivity to endocrine therapy[52]. NCT that modestly increases the pCR rate is unlikely to improve prognosis[54]. In this study, the increased sensitivity to chemotherapy did not fully compensate for the poor baseline prognosis and low sensitivity to endocrine therapy.

An important question is whether the results obtained in the neoadjuvant setting can be applied to the adjuvant setting. One limitation in this regard is that the prognostic information provided by ^{18}F-FDG metabolism was obtained from a large number of patients with advanced clinical stage. Gene expression studies indicate that primary tumor samples and metastatic lymph node samples from the same individual are usually more similar to

each other than to other samples, suggesting that the molecular program of primary breast cancer is retained in nodal metastases[7]. Furthermore, multigene assays have a prognostic value that is independent from the presence or absence of lymph node involvement[55-58]. In this study, we showed that the prognostic information on DRFS provided by the ^{18}F-FDG signature was comparable to that obtained from validated multigene prognostic signatures. This suggests that ^{18}F-FDG PET can be used to identify a low-risk population in the adjuvant setting when prognostic gene expression data are not available.

In this study, we analyzed the ESR1 signature as one of the multigene prognostic signatures. ER genes are significant determinants of intrinsic subtypes and multigene expression assays, and the biologic behavior of ER-low-positive cancers may be more similar to that of ER-negative cancers[59]. The ESR1 signature in this analysis was independently associated with chemotherapy response and DRFS. The prognostic information provided by this signature was comparable to that of other multigene prognostic signatures. Negative progesterone receptor status in the Asan cohort was also associated with pCR. Although we failed to demonstrate a statistically significant additive benefit of a

combined ^{18}F-FDG and ESR1 signature, the differential distribution of ^{18}F-FDG and ESR1 signature scores across intrinsic subtypes support the potential for integrating proliferation and ER signatures. The biologic activity derived from ER gene expression can be assessed noninvasively by ^{18}F-FDG PET[60]. Tumors with low baseline ^{18}F-fluoroestradiol (^{18}F-FES) uptake are unlikely to respond to endocrine therapy[61,62], whereas they may derive a greater benefit from chemotherapy[63]. Further prospective studies are needed to test the ability of the combination of ^{18}F-FDG and ^{18}F-FES PET (proliferation and ESR1 signature) to assess prognosis and guide therapy selection[64].

The present study had several limitations. First, the outcome data collected from the Asan cohort was based on events that had already occurred at the start of the study. Certain important clinical and pathological features, including ER expression levels, progesterone receptor status, and Ki-67 expression, were not available from the public cohorts. However, the Asan cohort included all known predictors and potential confounders. The statistical methods were predetermined according to the primary objectives in the study protocol. The possibility of information or selection bias was minimal. Second, the ^{18}F-FDG signature

score in this study might not represent tumor ^{18}F-FDG metabolism. The association between the ^{18}F-FDG signature score and the maximum SUV was moderate. We included a data series profiled using Affymetrix microarray platforms, and not all genes included in multigene assays are present on Affymetrix platforms. Accordingly, the genomic prognostic signatures analyzed may not represent the actual multigene assays; instead, they are genomic predictors based on similar prediction rules and the same genes. Nonetheless, the analysis indicated that all multigene signatures that are currently recommended by appropriate guidelines[48-50,65] were significantly associated with DRFS. Cellular proliferation requires the coordinated expression of hundreds of genes. Therefore, a large number of nominally different but equally good prognostic models can be built[66]. We believe that there might be redundancy in gene members, and our findings are generalizable. In addition, multigene assay results generated from fresh and formalin-fixed paraffin-embedded tissues may be equivalent[67]. The moderate association between the ^{18}F-FDG signature score and maximum SUV may be due to variability in ^{18}F-FDG uptake caused by the use of different PET scanners and imaging protocols without harmonization[68]. The SUV variability is an

important issue in multi-center clinical trials using PET/CT. Maximum SUV (single-pixel value) is reported to be highly affected by noise and leads to inaccurate results[69]. Although the Asan cohort performed SUV harmonization through matching recovery coefficient profiles, it is difficult to apply this method in other centers. SUV ratio (such as tumor to liver ratio) is thought to remove performance/reconstruction variability and may serve as a substitute for maximum SUV in multicenter trials. However, several studies have reported that the SUV ratio is also dependent on PET scanner and on image reconstruction method and thus needs to be harmonized[70]. Currently, international harmonization programs, such as European Association of Nuclear Medicine Research Ltd. program, is thought to be a feasible method for reducing SUV variability. If available, software solutions, such as EQ.PET, may also be an adequate method for PET quantification in multicenter trials[71,72].

Third, we did not establish the cut-off value of the maximum SUV to define poor and good prognosis for patients. The present results indicate that the harmonized maximum SUV of 4.1, which was the low tertile value, could be selected for a validation study based on the significance of the split in the survival curve.

Conclusion

Our data suggest that high tumor ^{18}F-FDG metabolism is associated with reduced DRFS and OS after adjusting for standard prognostic factors. ^{18}F-FDG metabolism provides prognostic information on DRFS that may be comparable to that obtained from validated multigene prognostic signatures. ^{18}F-FDG PET data may help classify ER-positive, ERBB2-negative patients into groups that would benefit from different therapeutic options.

References

1. Howlader N, Altekruse SF, Li CI, Chen VW, Clarke CA, Ries LA, et al: US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. J Natl Cancer Inst, 2014;106

2. Early Breast Cancer Trialists' Collaborative G, Davies C, Godwin J, Gray R, Clarke M, Cutter D, et al: Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. Lancet, 2011;378:771-84

3. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al: Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A, 2001;98:10869-74

4. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thurlimann B, Senn HJ, et al: Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. Ann Oncol, 2011;22:1736-47

5. Early Breast Cancer Trialists' Collaborative G: Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet, 2005;365:1687-717

6. Saura C, Hlauschek D, Oliveira M, Zardavas D, Jallitsch-Halper A, de la Pena L, et al: Neoadjuvant letrozole plus taselisib versus letrozole plus placebo in postmenopausal women with oestrogen receptor-positive, HER2-negative, early-stage breast cancer (LORELEI): a multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. Lancet Oncol, 2019;20:1226-1238

7. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al: Molecular portraits of human breast tumours. Nature, 2000;406:747-52

8. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al: A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med, 2004;351:2817-26

9. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, et al:

Gene expression profiling predicts clinical outcome of breast cancer. Nature, 2002;415:530-6

10. Wallden B, Storhoff J, Nielsen T, Dowidar N, Schaper C, Ferree S, et al: Development and verification of the PAM50-based Prosigna breast cancer gene signature assay. BMC Med Genomics, 2015;8:54

11. Denkert C, Kronenwett R, Schlake W, Bohmann K, Penzel R, Weber KE, et al: Decentral gene expression analysis for ER+/Her2- breast cancer: results of a proficiency testing program for the EndoPredict assay. Virchows Arch, 2012;460:251-9

12. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, et al: Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst, 2006;98:262-72

13. Carter SL, Eklund AC, Kohane IS, Harris LN, Szallasi Z: A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. Nat Genet, 2006;38:1043-8

14. Saal LH, Johansson P, Holm K, Gruvberger-Saal SK, She QB, Maurer M, et al: Poor prognosis in carcinoma is associated with a gene expression signature of aberrant PTEN tumor suppressor pathway activity. Proc Natl Acad Sci U S A, 2007;104:7564-9

15. Desmedt C, Haibe-Kains B, Wirapati P, Buyse M, Larsimont D, Bontempi G, et al: Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. Clin Cancer Res, 2008;14:5158-65

16. Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, et al: Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. Breast Cancer Res, 2008;10:R65

17. Ahn SG, Lee JH, Lee HW, Jeon TJ, Ryu YH, Kim KM, et al: Comparison of standardized uptake value of 18F-FDG-PET-CT with 21-gene recurrence score in estrogen receptor-positive, HER2-negative breast cancer. PLoS One, 2017;12:e0175048

18. Higuchi T, Nishimukai A, Ozawa H, Fujimoto Y, Yanai A, Miyagawa Y, et al: Prognostic significance of preoperative (18)F-FDG PET/CT for breast cancer subtypes. Breast, 2016;30:5-12

19. Palaskas N, Larson SM, Schultz N, Komisopoulou E, Wong J, Rohle D,

et al: 18F-fluorodeoxy-glucose positron emission tomography marks MYC-overexpressing human basal-like breast cancers. Cancer Res, 2011;71:5164-74

20. Farmer P, Bonnefoi H, Anderle P, Cameron D, Wirapati P, Becette V, et al: A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. Nat Med, 2009;15:68-74

21. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, et al: Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol, 2009;27:1160-7

22. Lee JK, Coutant C, Kim YC, Qi Y, Theodorescu D, Symmans WF, et al: Prospective comparison of clinical and genomic multivariate predictors of response to neoadjuvant chemotherapy in breast cancer. Clin Cancer Res, 2010;16:711-8

23. Juul N, Szallasi Z, Eklund AC, Li Q, Burrell RA, Gerlinger M, et al: Assessment of an RNA interference screen-derived mitotic and ceramide pathway metagene as a predictor of response to neoadjuvant paclitaxel for primary triple-negative breast cancer: a retrospective analysis of five clinical trials. Lancet Oncol, 2010;11:358-65

24. Iwamoto T, Bianchini G, Booser D, Qi Y, Coutant C, Shiang CY, et al: Gene pathways associated with prognosis and chemotherapy sensitivity in molecular subtypes of breast cancer. J Natl Cancer Inst, 2011;103:264-72

25. Hatzis C, Pusztai L, Valero V, Booser DJ, Esserman L, Lluch A, et al: A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. JAMA, 2011;305:1873-81

26. Horak CE, Pusztai L, Xing G, Trifan OC, Saura C, Tseng LM, et al: Biomarker analysis of neoadjuvant doxorubicin/cyclophosphamide followed by ixabepilone or Paclitaxel in early-stage breast cancer. Clin Cancer Res, 2013;19:1587-95

27. Stec J, Wang J, Coombes K, Ayers M, Hoersch S, Gold DL, et al: Comparison of the predictive accuracy of DNA array-based multigene classifiers across cDNA arrays and Affymetrix GeneChips. J Mol Diagn, 2005;7:357-67

28. Piccolo SR, Sun Y, Campbell JD, Lenburg ME, Bild AH, Johnson WE: A single-sample microarray normalization method to facilitate personalized-medicine workflows. Genomics, 2012;100:337-44

29. de Leeuw WC, Rauwerda H, Jonker MJ, Breit TM: Salvaging Affymetrix probes after probe-level re-annotation. BMC Res Notes, 2008;1:66

30. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD: The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics, 2012;28:882-3

31. Johnson WE, Li C, Rabinovic A: Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics, 2007;8:118-27

32. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science, 1999;286:531-7

33. Filipits M, Rudas M, Jakesz R, Dubsky P, Fitzal F, Singer CF, et al: A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. Clin Cancer Res, 2011;17:6012-20

34. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, et al: Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet, 2005;365:671-9

35. Gendoo DM, Ratanasirigulchai N, Schroder MS, Pare L, Parker JS, Prat A, et al: Genefu: an R/Bioconductor package for computation of gene expression-based signatures in breast cancer. Bioinformatics, 2016;32:1097-9

36. Chae SY, Son HJ, Lee DY, Shin E, Oh JS, Seo SY, et al: Comparison of diagnostic sensitivity of [(18)F]fluoroestradiol and [(18)F]fluorodeoxyglucose positron emission tomography/computed tomography for breast cancer recurrence in patients with a history of estrogen receptor-positive primary breast cancer. EJNMMI Res, 2020;10:54

37. Ha SC, Oh JS, Roh JL, Moon H, Kim JS, Cho KJ, et al: Pretreatment tumor SUVmax predicts disease-specific and overall survival in patients with head and neck soft tissue sarcoma. Eur J Nucl Med Mol Imaging, 2017;44:33-40

38. Lasnon C, Desmonts C, Quak E, Gervais R, Do P, Dubos-Arvis C, et al: Harmonizing SUVs in multicentre trials when using different generation PET systems: prospective validation in non-small cell lung cancer patients. Eur J Nucl Med Mol Imaging,

2013;40:985-96

39. Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, et al: FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging, 2015;42:328-54

40. Roepman P, de Koning E, van Leenen D, de Weger RA, Kummer JA, Slootweg PJ, et al: Dissection of a metastatic gene expression signature into distinct components. Genome Biol, 2006;7:R117

41. Soret M, Bacharach SL, Buvat I: Partial-volume effect in PET tumor imaging. J Nucl Med, 2007;48:932-45

42. McCarthy DJ, Chen Y, Smyth GK: Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res, 2012;40:4288-97

43. Hudis CA, Barlow WE, Costantino JP, Gray RJ, Pritchard KI, Chapman JA, et al: Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: the STEEP system. J Clin Oncol, 2007;25:2127-32

44. Miyake T, Nakayama T, Naoi Y, Yamamoto N, Otani Y, Kim SJ, et al: GSTP1 expression predicts poor pathological complete response to neoadjuvant chemotherapy in ER-negative breast cancer. Cancer Sci, 2012;103:913-20

45. Jin S, Kim SB, Ahn JH, Jung KH, Ahn SH, Son BH, et al: 18 F-fluorodeoxyglucose uptake predicts pathological complete response after neoadjuvant chemotherapy for breast cancer: a retrospective cohort study. J Surg Oncol, 2013;107:180-7

46. Yoo C, Ahn JH, Jung KH, Kim SB, Kim HH, Shin HJ, et al: Impact of immunohistochemistry-based molecular subtype on chemosensitivity and survival in patients with breast cancer following neoadjuvant chemotherapy. J Breast Cancer, 2012;15:203-10

47. Kwon Y, Ro J, Kang HS, Kim SK, Hong EK, Khang SK, et al: Clinicopathological parameters and biological markers predicting non-sentinel node metastasis in sentinel node-positive breast cancer patients. Oncol Rep, 2011;25:1063-71

48. Cardoso F, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rubio IT, et al: Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol, 2019;30:1194-1220

49. Andre F, Ismaila N, Henry NL, Somerfield MR, Bast RC, Barlow W, et al: Use of Biomarkers to Guide Decisions on Adjuvant Systemic Therapy for Women With Early-Stage Invasive Breast Cancer: ASCO Clinical Practice Guideline Update-Integration of Results From TAILORx. J Clin Oncol, 2019;37:1956-1964

50. Henry NL, Somerfield MR, Abramson VG, Ismaila N, Allison KH, Anders CK, et al: Role of Patient and Disease Factors in Adjuvant Systemic Therapy Decision Making for Early-Stage, Operable Breast Cancer: Update of the ASCO Endorsement of the Cancer Care Ontario Guideline. J Clin Oncol, 2019;37:1965-1977

51. White IR, Royston P, Wood AM: Multiple imputation using chained equations: Issues and guidance for practice. Stat Med, 2011;30:377-99

52. Liedtke C, Hatzis C, Symmans WF, Desmedt C, Haibe-Kains B, Valero V, et al: Genomic grade index is associated with response to chemotherapy in patients with breast cancer. J Clin Oncol, 2009;27:3185-91

53. Iwamoto T, Lee JS, Bianchini G, Hubbard RE, Young E, Matsuoka J, et al: First generation prognostic gene signatures for breast cancer predict both survival and chemotherapy sensitivity and identify overlapping patient populations. Breast Cancer Res Treat, 2011;130:155-64

54. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, et al: Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. Lancet, 2014;384:164-72

55. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, et al: Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol, 2010;11:55-65

56. Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delaloge S, et al: 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. N Engl J Med, 2016;375:717-29

57. Laenkholm AV, Jensen MB, Eriksen JO, Rasmussen BB, Knoop AS, Buckingham W, et al: PAM50 Risk of Recurrence Score Predicts 10-Year Distant Recurrence in a Comprehensive Danish Cohort of Postmenopausal Women Allocated to 5 Years of

Endocrine Therapy for Hormone Receptor-Positive Early Breast Cancer. J Clin Oncol, 2018;36:735-740

58. Sestak I, Buus R, Cuzick J, Dubsky P, Kronenwett R, Denkert C, et al: Comparison of the Performance of 6 Prognostic Signatures for Estrogen Receptor-Positive Breast Cancer: A Secondary Analysis of a Randomized Clinical Trial. JAMA Oncol, 2018;4:545-553

59. National Comprehensive Caner Network. NCCN clinical practice guidelines in oncology: breast cancer, version 4.2020. 2020;https://www.nccn.org/professionals/physician_gls/pdf/breast.pdf:Assessed June 1, 2020

60. Chae SY, Ahn SH, Kim SB, Han S, Lee SH, Oh SJ, et al: Diagnostic accuracy and safety of 16alpha-[(18)F]fluoro-17beta-oestradiol PET-CT for the assessment of oestrogen receptor status in recurrent or metastatic lesions in patients with breast cancer: a prospective cohort study. Lancet Oncol, 2019;20:546-555

61. Mortimer JE, Dehdashti F, Siegel BA, Trinkaus K, Katzenellenbogen JA, Welch MJ: Metabolic flare: indicator of hormone responsiveness in advanced breast cancer. J Clin Oncol, 2001;19:2797-803

62. Linden HM, Stekhova SA, Link JM, Gralow JR, Livingston RB, Ellis GK, et al: Quantitative fluoroestradiol positron emission tomography imaging predicts response to endocrine treatment in breast cancer. J Clin Oncol, 2006;24:2793-9

63. Chae SY, Kim SB, Ahn SH, Kim HO, Yoon DH, Ahn JH, et al: A Randomized Feasibility Study of (18)F-Fluoroestradiol PET to Predict Pathologic Response to Neoadjuvant Therapy in Estrogen Receptor-Rich Postmenopausal Breast Cancer. J Nucl Med, 2017;58:563-568

64. Kurland BF, Peterson LM, Lee JH, Schubert EK, Currin ER, Link JM, et al: Estrogen Receptor Binding (18F-FES PET) and Glycolytic Activity (18F-FDG PET) Predict Progression-Free Survival on Endocrine Therapy in Patients with ER+ Breast Cancer. Clin Cancer Res, 2017;23:407-415

65. Duffy MJ, Harbeck N, Nap M, Molina R, Nicolini A, Senkus E, et al: Clinical use of biomarkers in breast cancer: Updated guidelines from the European Group on Tumor Markers (EGTM). Eur J Cancer, 2017;75:284-298

66. Gyorffy B, Hatzis C, Sanft T, Hofstatter E, Aktas B, Pusztai L: Multigene prognostic tests in breast cancer: past, present, future. Breast Cancer Res, 2015;17:11

67. Beumer I, Witteveen A, Delahaye L, Wehkamp D, Snel M, Dreezen C, et al: Equivalence of MammaPrint array types in clinical trials and diagnostics. Breast Cancer Res Treat, 2016;156:279-87

68. Boellaard R: Standards for PET image acquisition and quantitative data analysis. J Nucl Med, 2009;50 Suppl 1:11S-20S

69. Boellaard R, Krak NC, Hoekstra OS, Lammertsma AA: Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. J Nucl Med, 2004;45:1519-27

70. Quak E, Hovhannisyan N, Lasnon C, Fruchart C, Vilque JP, Musafiri D, et al: The importance of harmonizing interim positron emission tomography in non-Hodgkin lymphoma: focus on the Deauville criteria. Haematologica, 2014;99:e84-5

71. Aide N, Lasnon C, Veit-Haibach P, Sera T, Sattler B, Boellaard R: EANM/EARL harmonization strategies in PET quantification: from daily practice to multicentre oncological studies. Eur J Nucl Med Mol Imaging, 2017;44:17-31

72. Lasnon C, Salomon T, Desmonts C, Do P, Oulkhouir Y, Madelaine J, et al: Generating harmonized SUV within the EANM EARL accreditation program: software approach versus EARL-compliant reconstruction. Ann Nucl Med, 2017;31:125-134

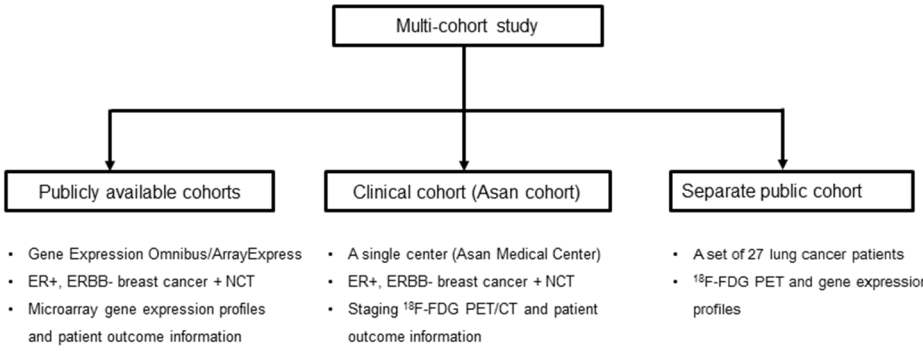


Fig. 1 Study design. The multi-cohort study composed of three cohort groups. ER = estrogen receptor; NCT = neoadjuvant chemotherapy

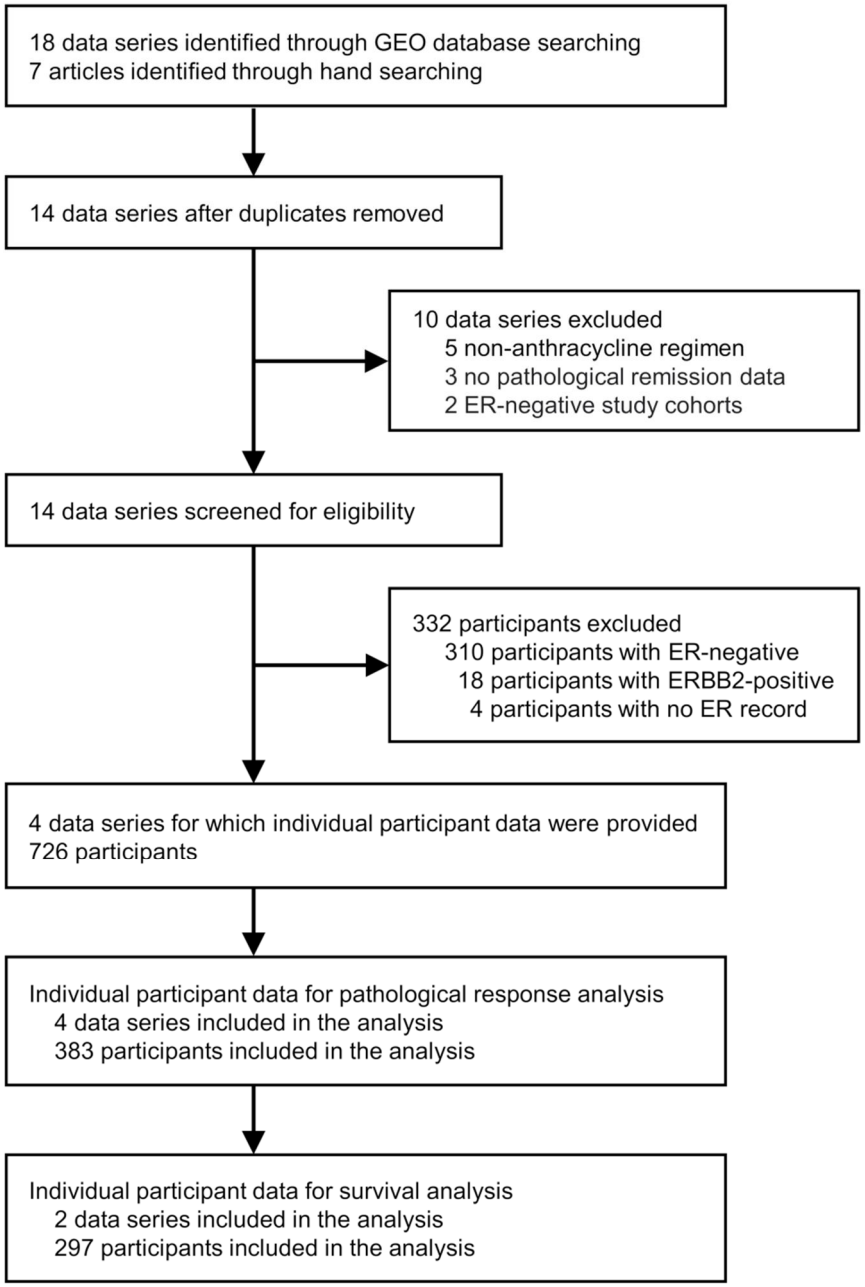


Fig. 2 Flow diagram for the public microarray datasets. The pooled analysis included 297 participants with available microarray data sets. GEO = Gene Expression Omnibus.

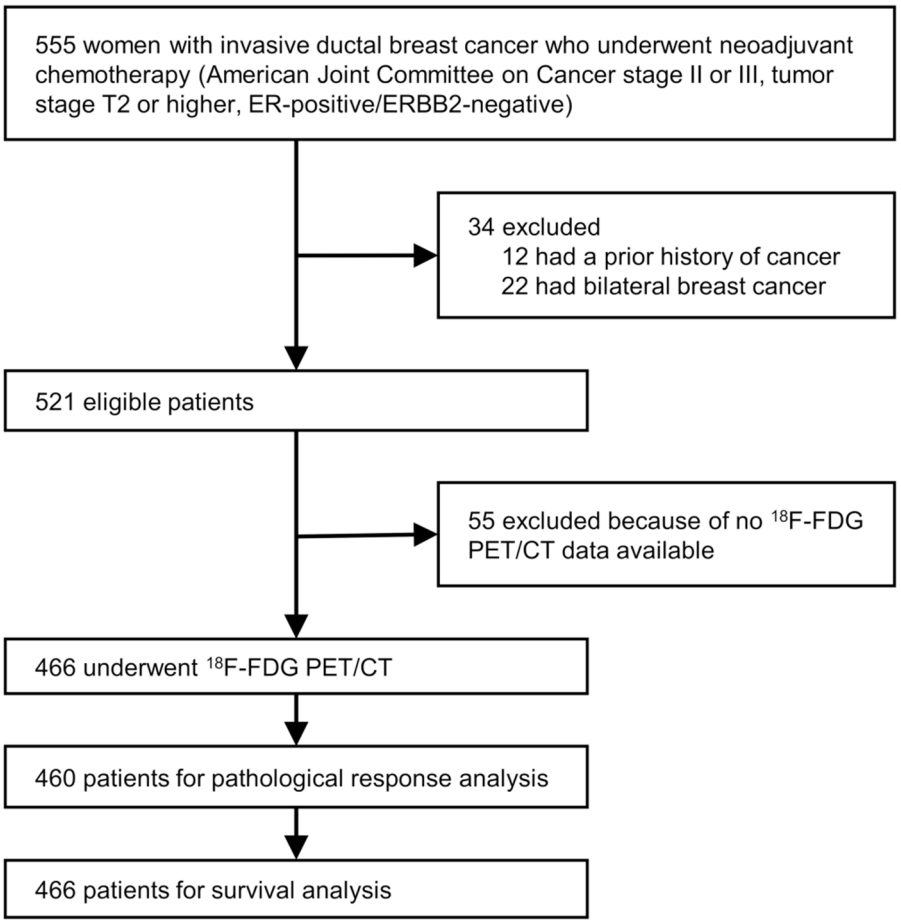


Fig. 3 Flow diagram for the Asan cohort. Pathological complete response and survival were investigated in the Asan Medical Center cohort of ER-positive, ERBB2-negative breast cancer patients.

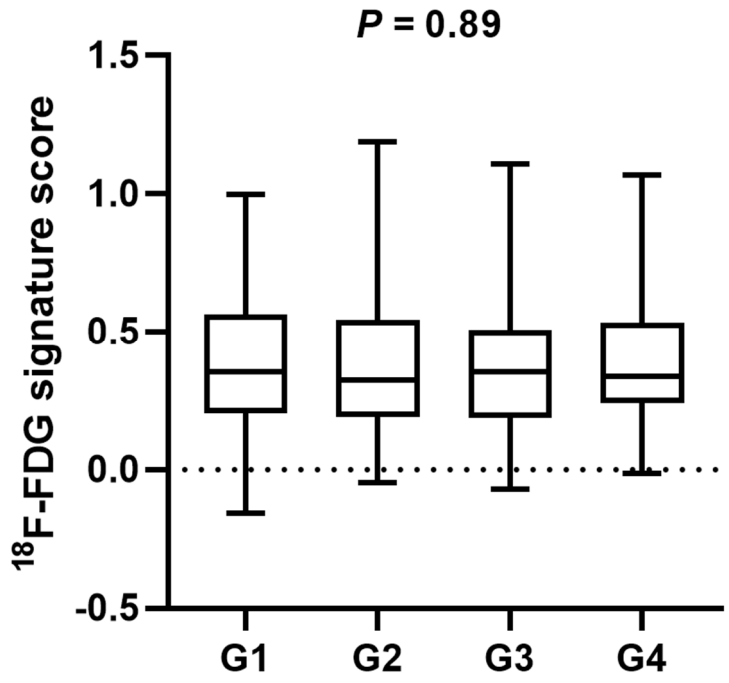


Fig. 4 ^{18}F-fluorodeoxyglucose (^{18}F-FDG) signature scores of four publicly available breast cancer data sets. The upper and lower whiskers represent the upper 25% and lower 25% of scores, respectively, with the minimum and maximum scores shown at the end. The box plot represents the middle 50% of scores. The median marks the mid-point of the data shown by the line that divides the box into two parts. G1, GSE22093; G2, GSE25055; G3, GSE25065; G4, GSE32646.

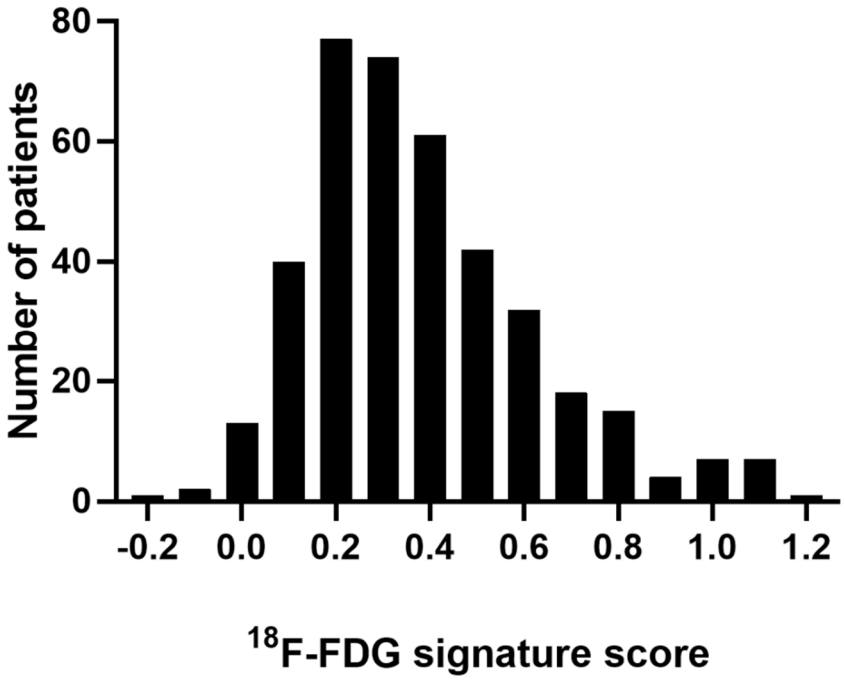


Fig. 5 Distribution of ^{18}F-fluorodeoxyglucose (^{18}F-FDG) signature scores from the publicly available Gene Expression Omnibus repository.

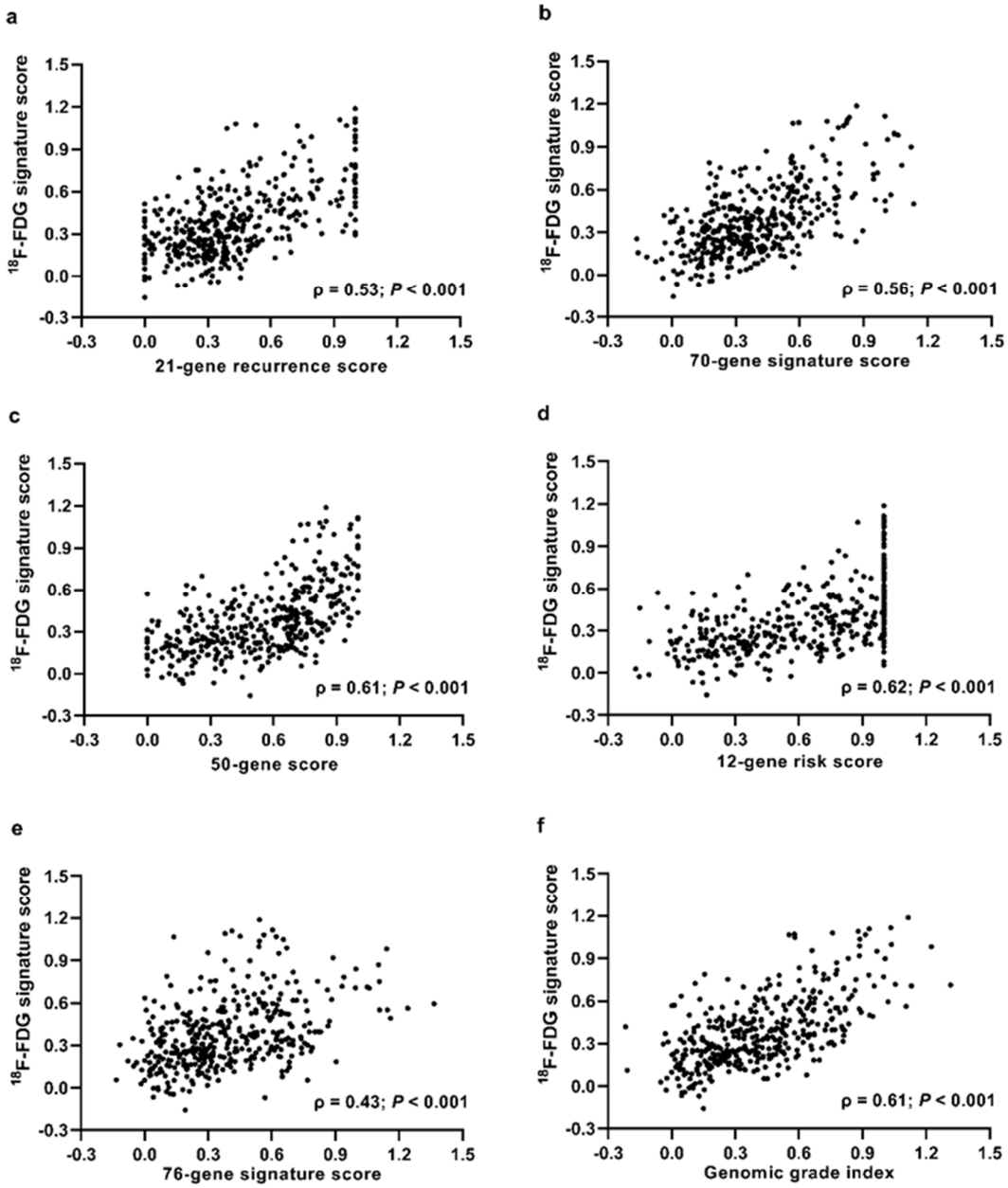


Fig. 6 Association of ^{18}F-fluorodeoxyglucose (^{18}F-FDG) scores with multigene prognostic signatures.

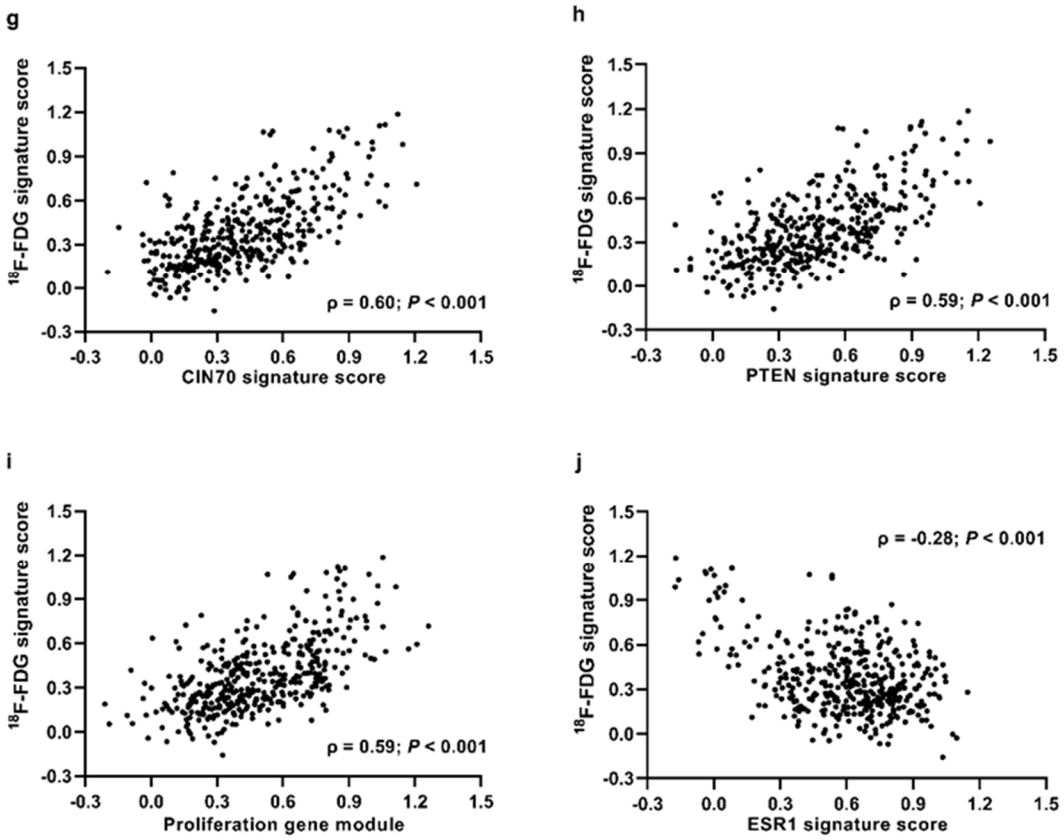


Fig. 6 Association of ^{18}F-fluorodeoxyglucose (^{18}F-FDG) scores with multigene prognostic signatures (continued). The ^{18}F-FDG signature score was plotted against the 21-gene recurrence score (a), 70-gene signature score (b), 50-gene score (c), 12-gene risk score (d), 76-gene signature score (e), genomic grade index (f), CIN70 signature score (g), PTEN signature score (h), proliferation gene module (i), and ESR1 signature score (j).

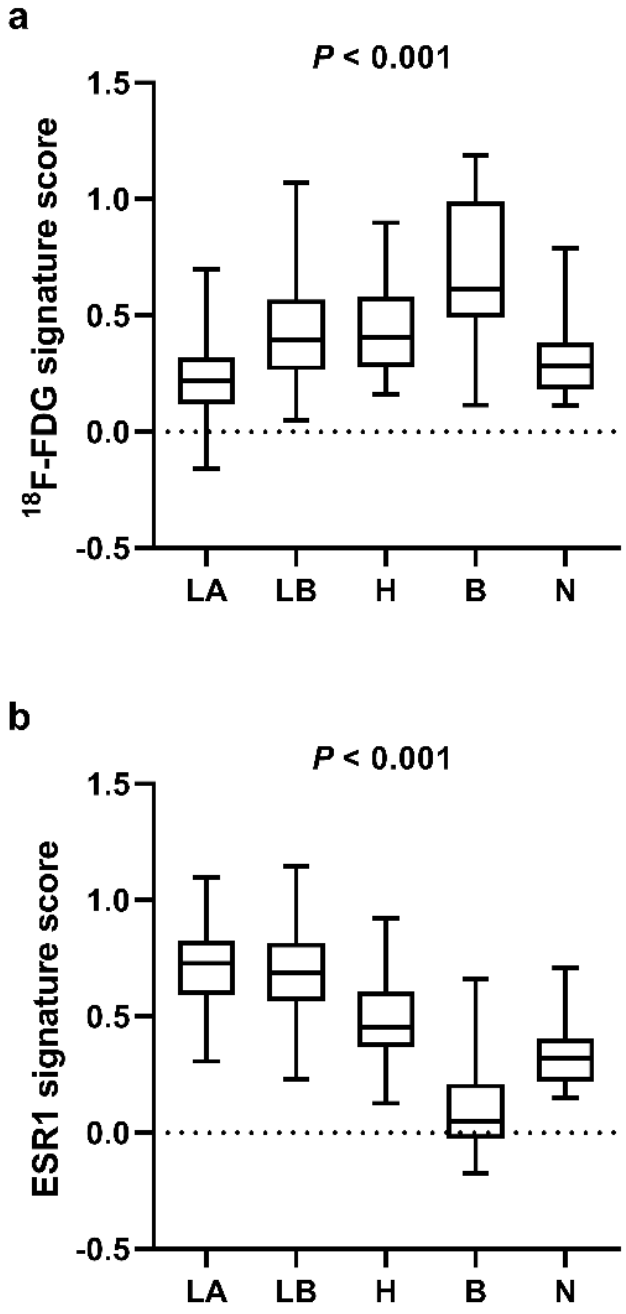


Fig. 7 ^{18}F-fluorodeoxyglucose (^{18}F-FDG) and estrogen receptor 1 (ESR1) expression signature scores of the intrinsic molecular subtypes. Breast cancers from public cohorts were analyzed. The upper and lower whiskers represent the upper 25% and lower 25% of

scores, respectively, with the minimum and maximum scores shown at the end. The box plot represents the middle 50% of scores. The median marks the mid-point of the data indicated by the line that divides the box into two parts. Five intrinsic molecular subtypes were classified by a 50-gene single sample predictor. (a) ^{18}F-FDG signature scores differed significantly according to the intrinsic molecular subtype ($P < .001$). The luminal A and normal-like subtypes had low ^{18}F-FDG signature scores, whereas the luminal B, HER2-enriched, and basal-like subtypes had significantly higher ^{18}F-FDG signature scores than the luminal A subtype ($P < .001$). (b) ESR1 signature scores differed significantly according to the intrinsic molecular subtype ($P < .001$). The luminal A and B subtypes had significantly higher ESR1 scores than other subtypes ($P < .001$). LA, luminal A; LB, luminal B; H, HER2-enriched; B, basal-like; N, normal-like.

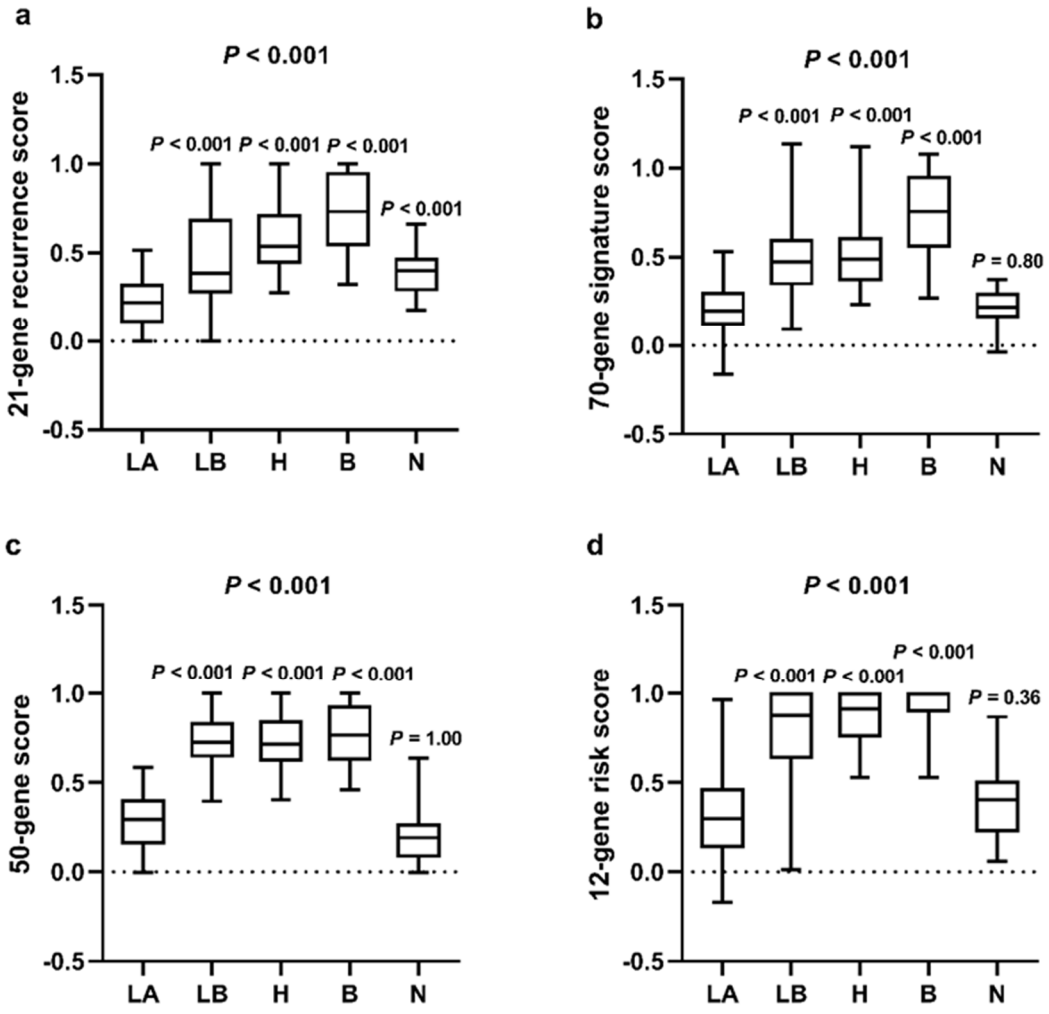


Fig. 8 Multigene prognostic signature scores of the intrinsic molecular subtypes.

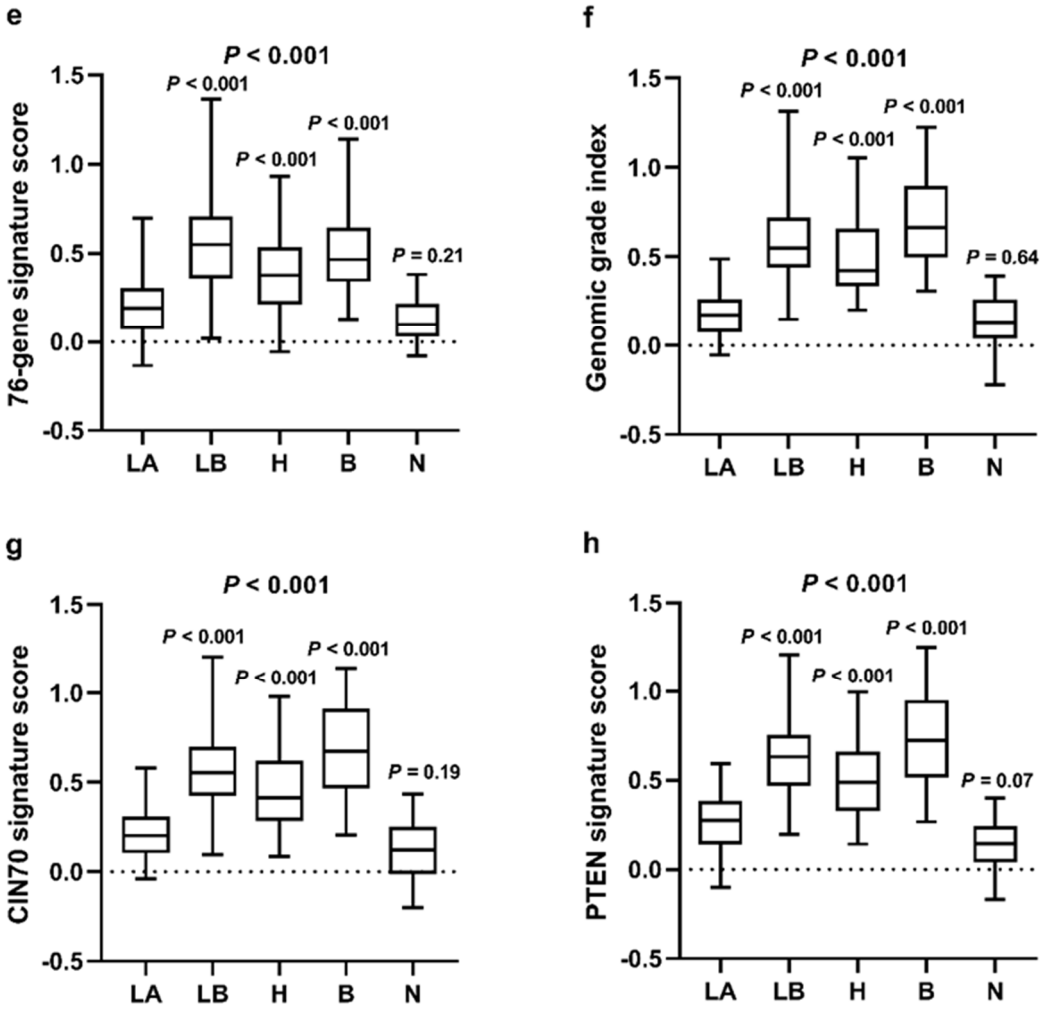


Fig. 8 Multigene prognostic signature scores of the intrinsic molecular subtypes (continued).

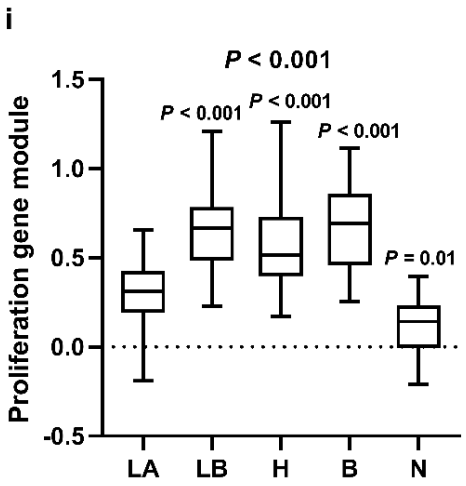


Fig. 8 Multigene prognostic signature scores of the intrinsic molecular subtypes (continued). Breast cancers from the publicly available Gene Expression Omnibus data sets were analyzed. The upper and lower whiskers represent the upper 25% and lower 25% of scores, respectively, with the minimum and maximum scores shown at the end. The box plot represents the middle 50% of scores. The median marks the mid-point of the data indicated by the line that divides the box into two parts. Multigene prognostic signature scores included the 21-gene recurrence score (a), 70-gene signature score (b), 50-gene score (c), 12-gene risk score (d), 76-gene signature score (e), genomic grade index (f), CIN70 signature score (g), PTEN signature score (h), and proliferation gene module (i). Five intrinsic molecular subtypes were classified by a 50-gene single sample predictor. LA, luminal A; LB,

luminal B; H, HER2-enriched; B, basal-like; N, normal-like. P values calculated with the Kruskal-Wallis test with post hoc Dunn's test between luminal A and other subtypes are shown.

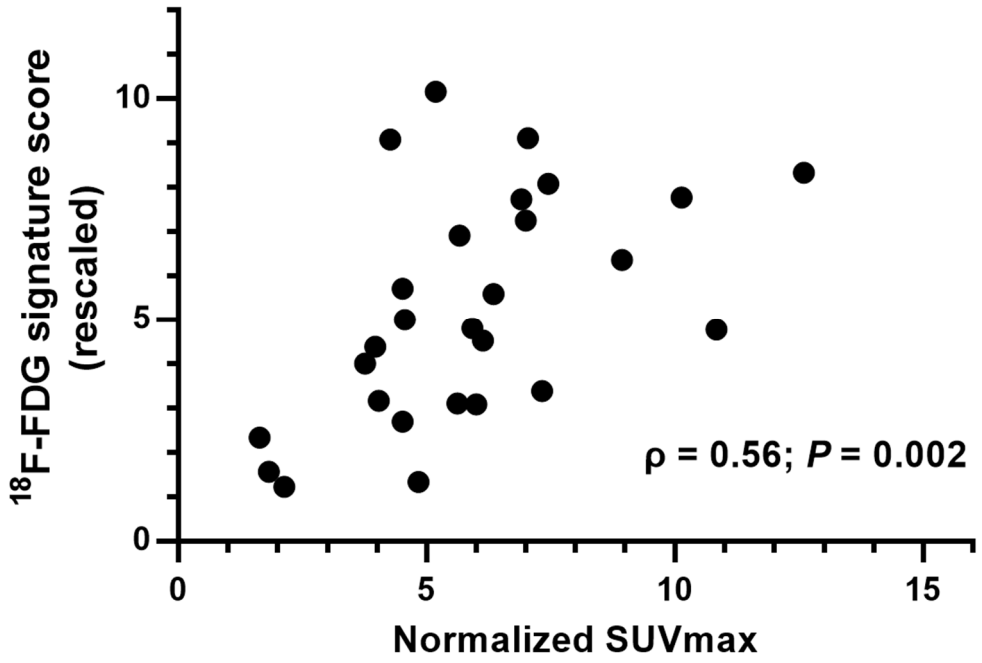


Fig. 9 Association between the ^{18}F-fluorodeoxyglucose (^{18}F-FDG) signature score and normalized maximum standardized uptake value (SUV) of ^{18}F-FDG in patients with lung cancer.

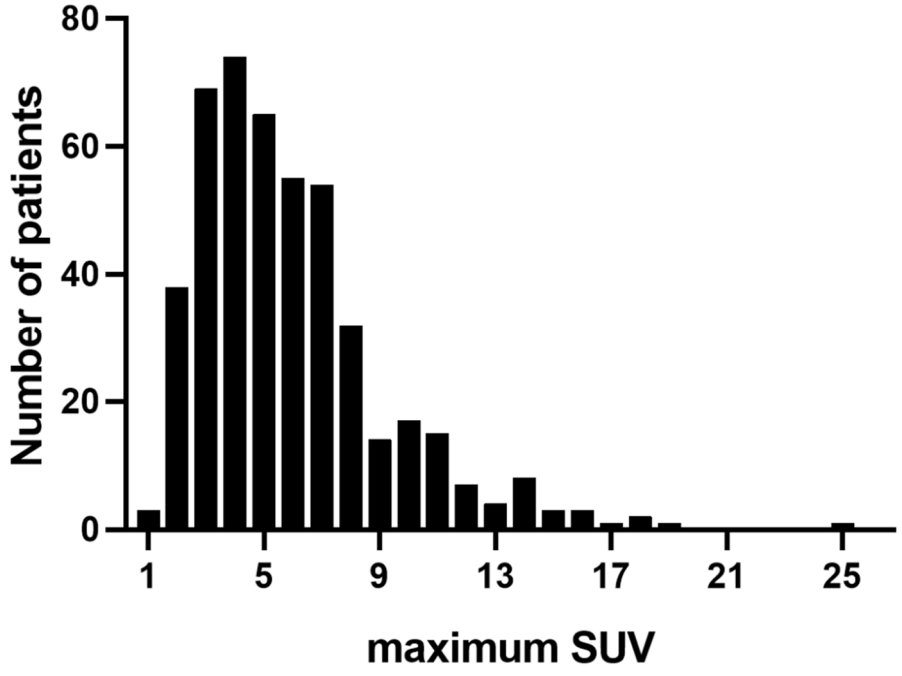


Fig. 10 Distribution of maximum standardized uptake value (SUV) of ^{18}F-fluorodeoxyglucose (^{18}F-FDG) positron emission tomography/computed tomography (PET/CT) in patients with breast cancer.

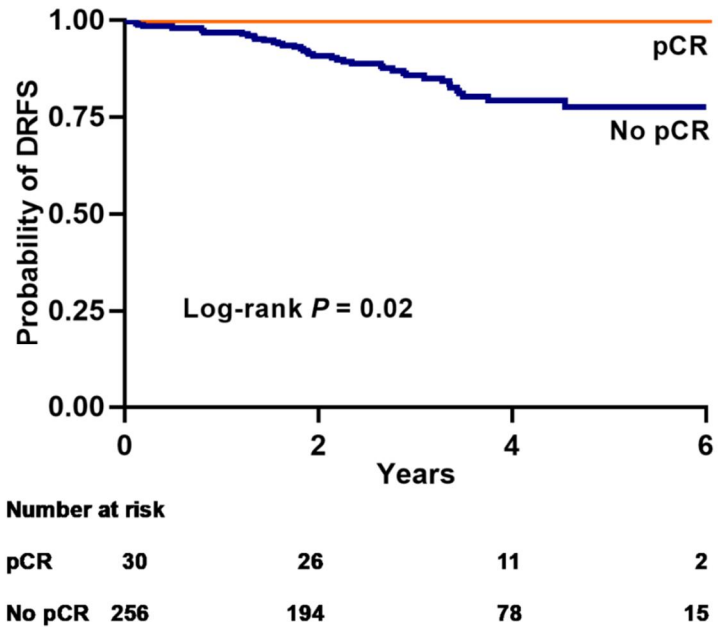


Fig. 11 Kaplan-Meier estimates of distant relapse-free survival (DRFS) from publicly available cohorts according to pathological complete response (pCR) after neoadjuvant chemotherapy. Distant relapse occurred in none of 30 (0%) patients who achieved pCR, whereas it occurred in 40 of 256 (15.6%) patients who did not achieve pCR. Patients with pCR had significantly longer DRFS (P = .02).

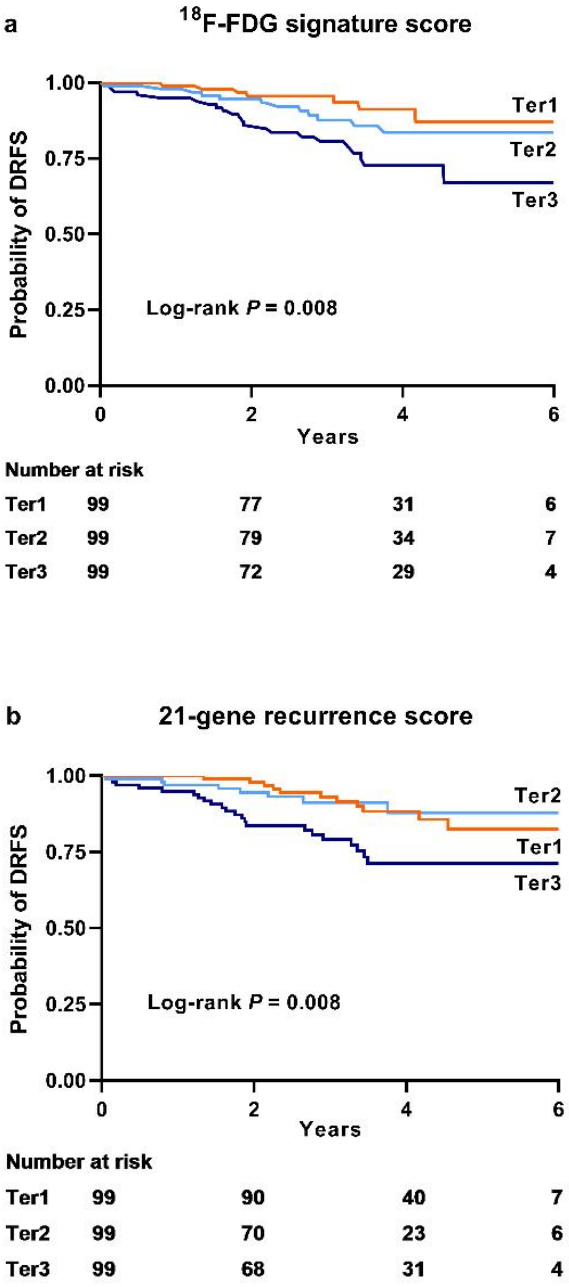


Fig. 12 Prognostic distant relapse-free survival (DRFS) analysis according to tertiles of multigene signature scores.

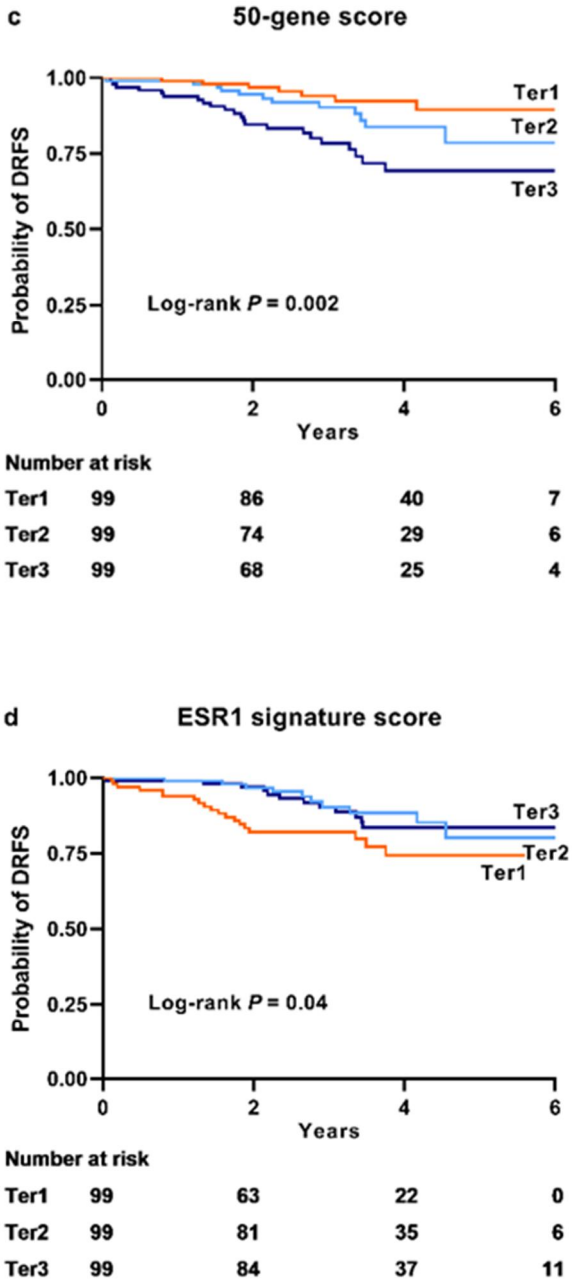


Fig. 12 Prognostic distant relapse-free survival (DRFS) analysis according to tertiles of multigene signature scores (continued). (a) ^{18}F-fluorodeoxyglucose (^{18}F-FDG) signature score; (b) 21-gene recurrence score; (c) 50-gene score; (d) estrogen receptor 1 (ESR1)

signature score. Multigene prognostic signatures, except the 76-gene, CIN70, and proliferation gene module, were significantly associated with DRFS.

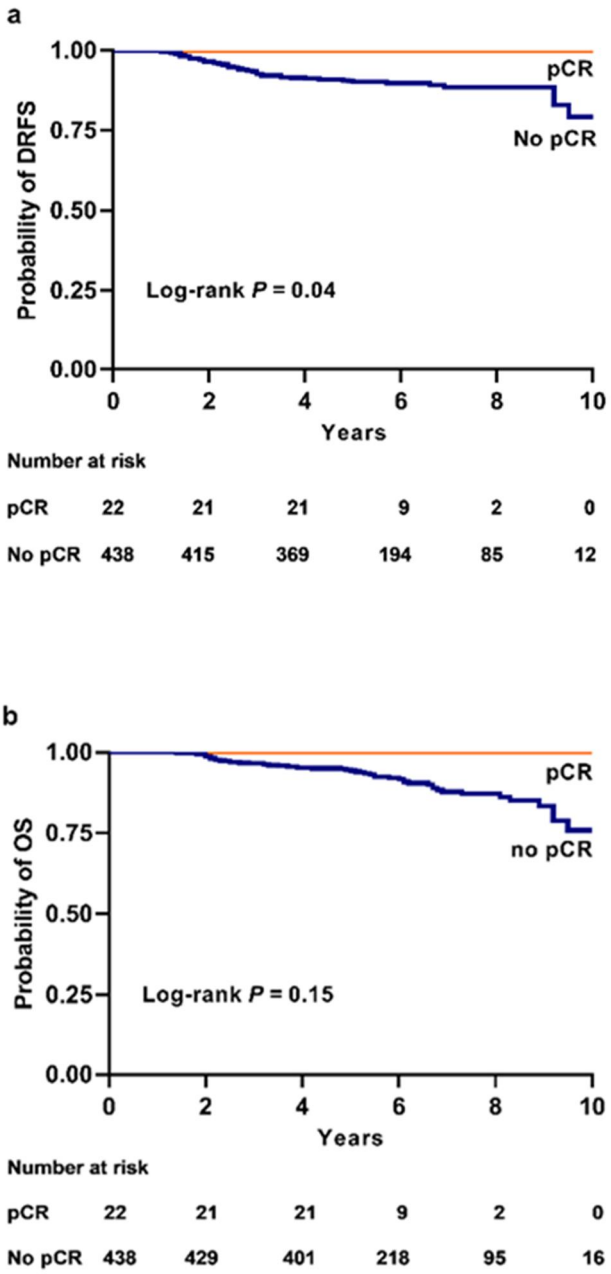


Fig. 13 Kaplan-Meier estimates of distant relapse-free survival (DRFS) and overall survival (OS) of breast cancer patients according to pathological complete response (pCR) after neoadjuvant chemotherapy. (a) Distant relapse occurred in none of 22 patients

who achieved pCR. However, among 438 patients who did not achieve pCR, distant relapse occurred in 81 (18.5%) patients, and the difference between the groups was significant (*P* = .04). (b) No death was observed among 22 patients who achieved pCR, whereas 49 (11.1%) patients who did not achieve pCR died during the follow-up period.

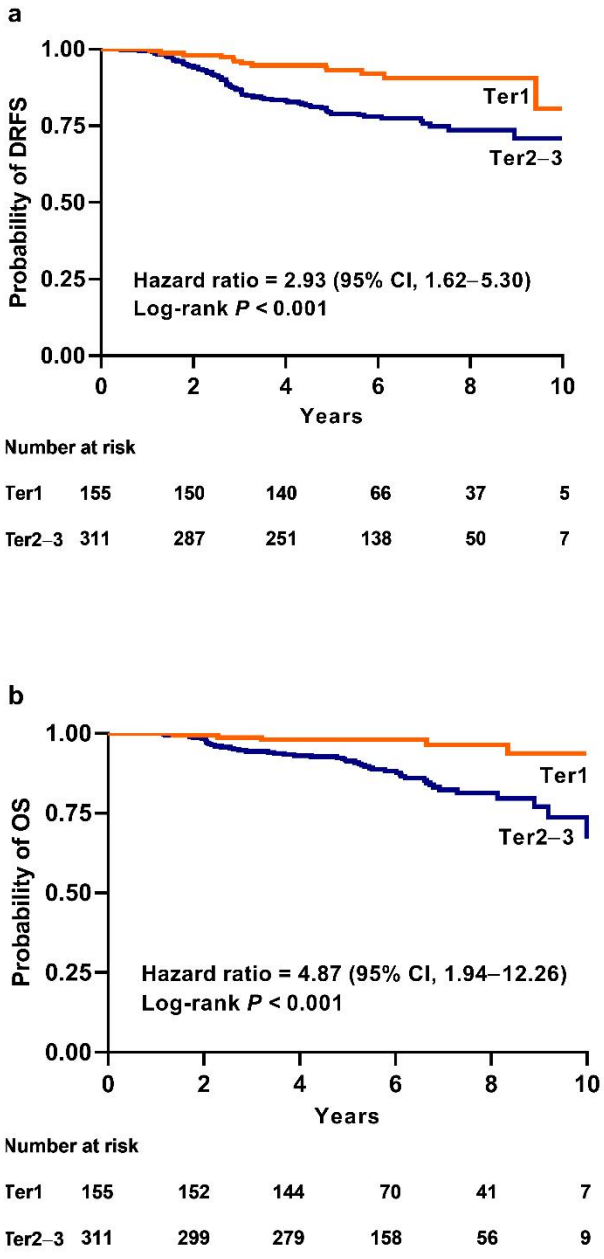


Figure 14 Kaplan-Meier curves of distant relapse-free survival and overall survival by tertiles of the maximum standardized uptake value of ^{18}F-fluorodeoxyglucose. The middle and high tertile categories were combined into one high-risk group because their

outcome was significantly different from that of the low tertile. (a) The middle and high tertiles of SUV were prognostic for DRFS (hazard ratio = 2.93, 95% confidence interval [CI] = 1.62–5.30; $P < 0.001$). The 8-year DRFS rates were 90.7% (95% CI, 85.5–96.1%) for those in the low tertile of maximum SUV vs. 73.7% (95% CI = 68.0–79.8%) for those in the middle and high tertiles of maximum SUV. (b) The middle and high tertiles of SUV were prognostic for OS (hazard ratio = 4.87, 95% CI = 1.94–12.26; $P < 0.001$). The 8-year OS rates were 96.4% (95% CI = 92.6–100%) for those in the low tertile of maximum SUV vs. 81.3% (95% CI = 76.0–87.0%) for those in the middle and high tertiles of maximum SUV. DRFS = distant relapse-free survival; OS = overall survival; SUV = standardized uptake value; Ter1 = low tertile; Ter2–3 = middle and high tertiles.

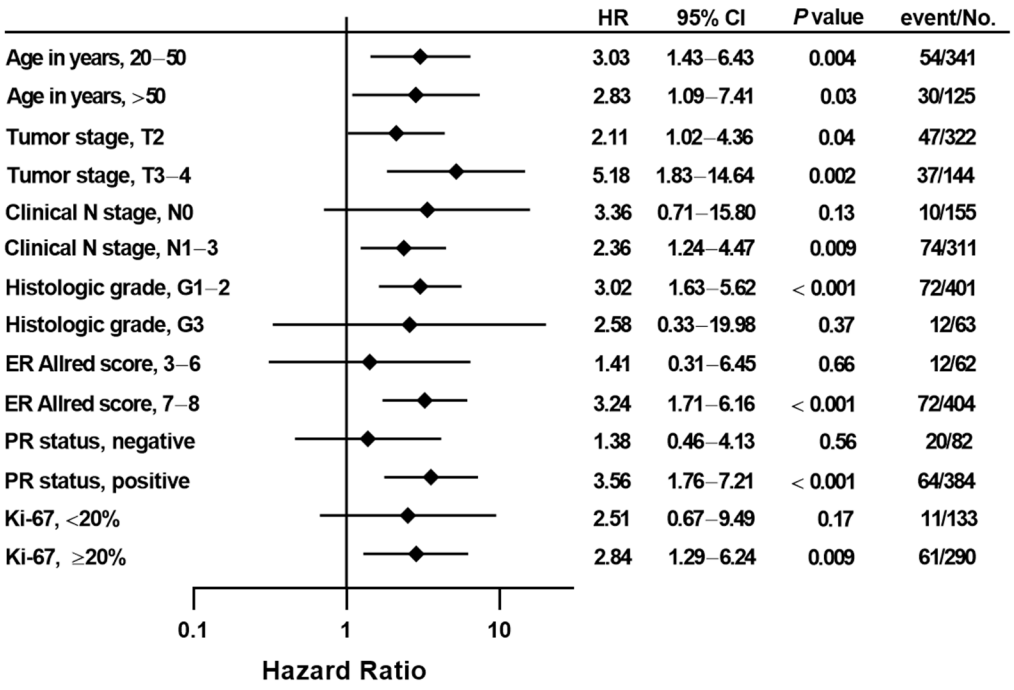


Fig. 15 Extended Cox proportional hazard analyses of maximum standardized uptake value for distant relapse-free survival. Subgroup analysis evaluating the association between the maximum standardized uptake value and distant relapse-free survival according to clinical and pathological characteristics. The maximum standardized uptake value was dichotomized based on the cut-off value of the low tertile (the low tertile vs. the middle and high tertiles).

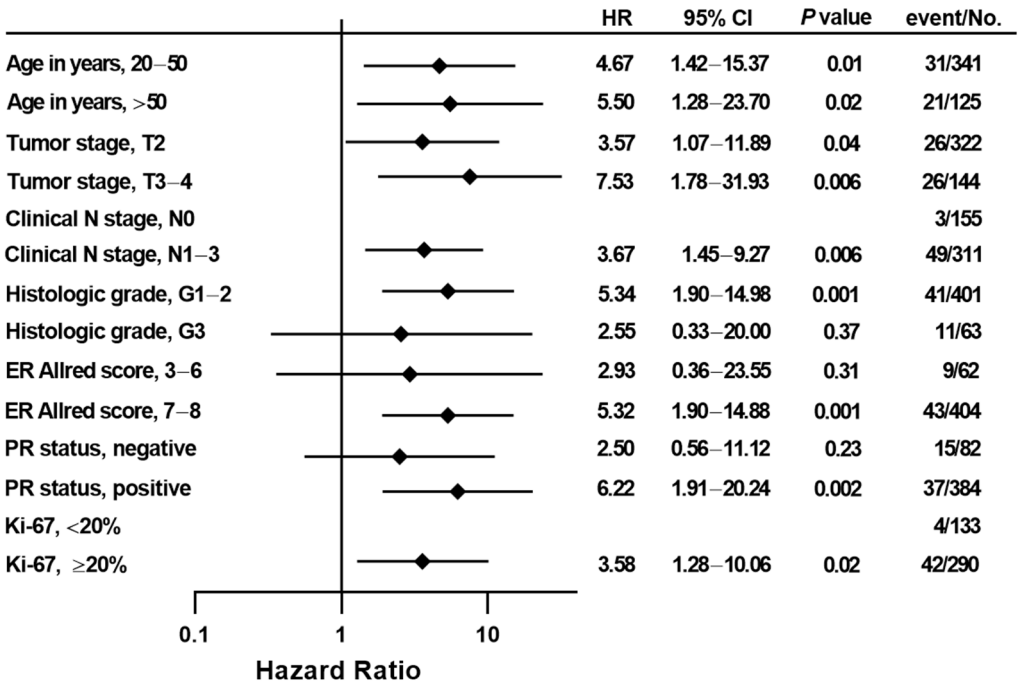


Fig. 16 Extended Cox proportional hazard analyses of maximum standardized uptake value for overall survival. Subgroup analysis evaluating the association between the maximum standardized uptake value and overall survival according to clinical and pathological characteristics. The maximum standardized uptake value was dichotomized based on the cut-off value of the low tertile (the low tertile vs. the middle and high tertile).

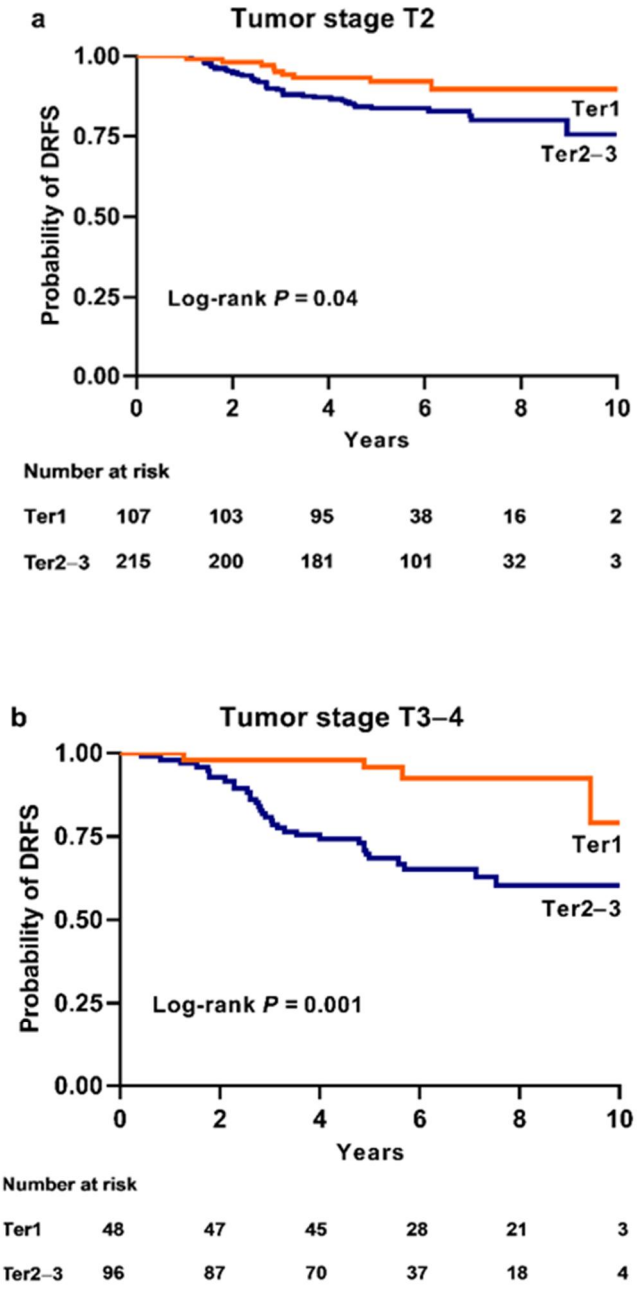


Fig. 17 Kaplan-Meier curves of distant relapse-free survival (DRFS) according to tertiles of the maximum standardized uptake value (SUV) of ^{18}F-fluorodeoxyglucose.

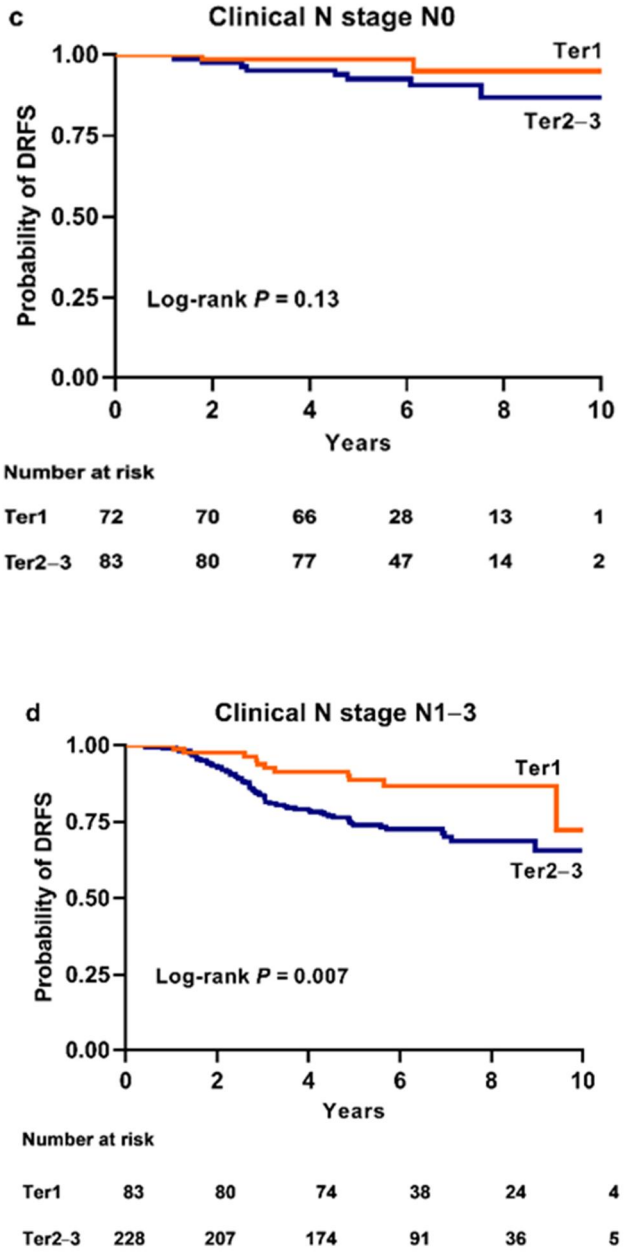


Fig. 17 Kaplan-Meier curves of distant relapse-free survival (DRFS) according to tertiles of the maximum standardized uptake value (SUV) of ^{18}F-fluorodeoxyglucose (continued). Subgroup analysis evaluating the association between the maximum SUV and

DRFS according to tumor stage (a: tumor stage T2; b: tumor stage T3–4) and clinical N stage (c: node-negative disease, N0; d: node-positive disease, N1–3). The maximum SUV was dichotomized based on the cut-off value of the low tertile. Ter1 indicates a subgroup of patients in the low tertile of SUV (1.36–4.14); Ter2-3 represents patients with the middle and high tertiles of SUV (4.14–25.06). The 8-year DRFS rates (with 95% CI) of patients in the low tertile of the maximum SUV vs. those in the middle, or the high tertile of SUVmax were (a) 89.7% (83.0–96.9%) vs. 80.1% (74.0–86.6%), (b) 92.4% (84.3–100.0%) vs. 60.3% (49.9–72.8%), (c) 94.9% (87.7–100%) vs. 86.8% (77.6–97.1%), and (d) 86.8% (79.4–94.9%) vs. 68.8% (62.0–76.3%).

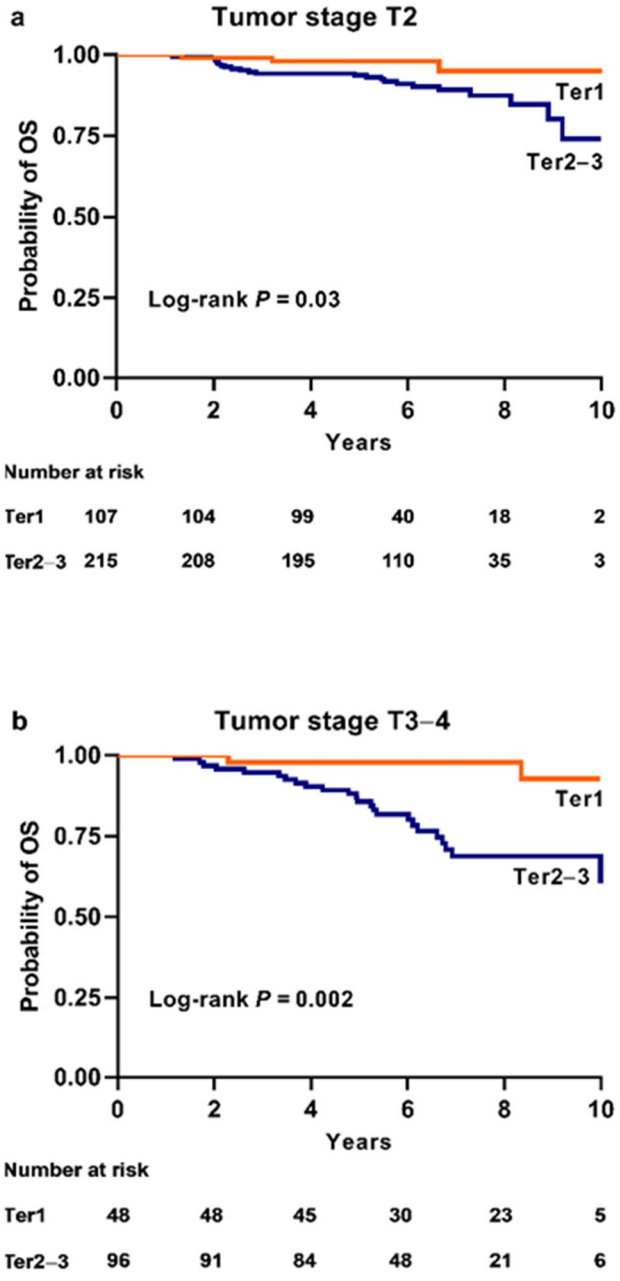


Fig. 18 Kaplan-Meier curves of overall survival (OS) according to tertiles of the maximum standardized uptake value (SUV) of ^{18}F-fluorodeoxyglucose.

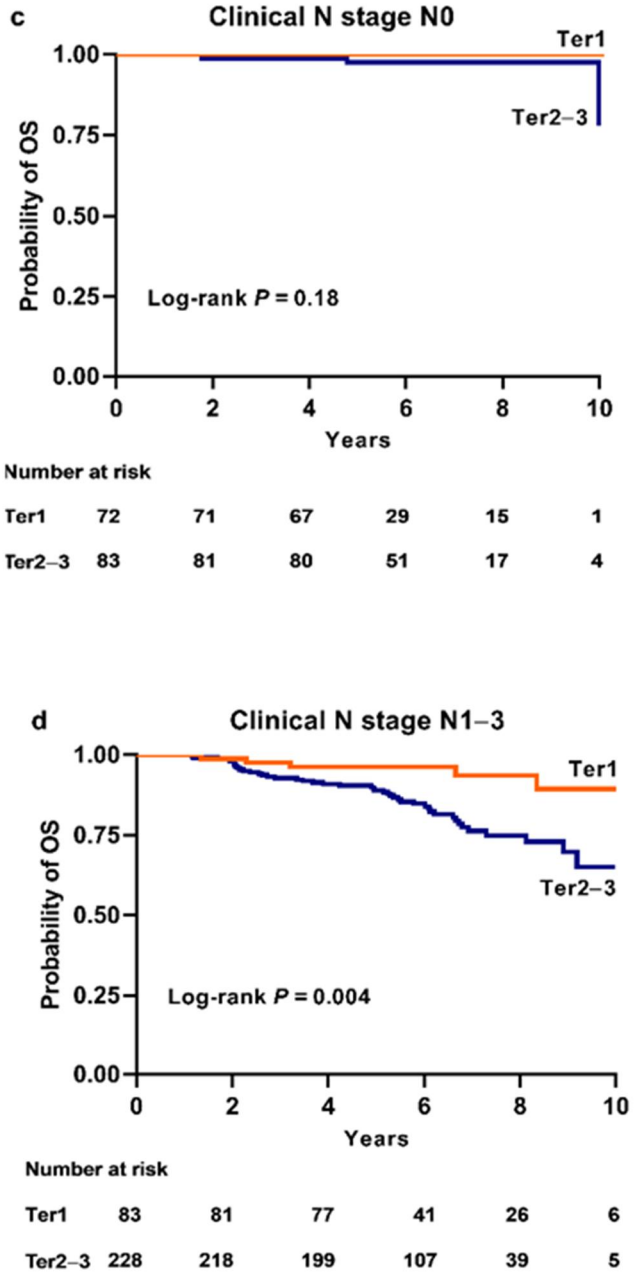


Fig. 18 Kaplan-Meier curves of overall survival (OS) according to tertiles of the maximum standardized uptake value (SUV) of ^{18}F-fluorodeoxyglucose (continued). Subgroup analysis evaluating the association between the maximum SUV and OS according to tumor stage (a: tumor stage T2; b: tumor stage T3–4) and clinical N stage (c: node-

negative disease, N0; d: node-positive disease, N1–3). The maximum SUV was dichotomized based on the cut-off value of the low tertile. Ter1 indicates a subgroup of patients in the low tertile of SUV (1.36–4.14); Ter2–3 represents patients in the middle and high tertiles of SUV (4.14–25.06). The 8-year OS rates (with 95% CI) of patients in the low tertile of the maximum SUV vs. those in the middle, or the high tertile of SUVmax were (a) 95.0% (88.8–100.0%) vs. 87.5% (81.9–93.4%), (b) 97.9% (94.0–100.0%) vs. 68.9% (58.6–81.0%), (c) 100.0% (100.0–100.0%) vs. 97.5% (94.2–100.0%), and (d) 93.6% (87.4–100.0%) vs. 74.9% (67.9–82.6%).

Table 1. Clinical and pathological characteristics of the public and Asan cohorts

Characteristics	Public data series from GEO					Asan cohort
	[a]MADCC IGR	[b]MADCC I-SPY-1	[c]MADCC LBJ, INEN GEICAM USO	[d]Miyake	Total	
Chemotherapy	FAC	TFAC/ACT	TFAC/FEC	TFEC		AC/AT/ACT
No. of patients	42	174	123	55	394	466
Age, years						
20–50	27	93 (53.4%)	65 (52.8%)	31 (56.4%)	216 (54.8%)	341 (73.2%)
>50	15	81 (46.6%)	58 (47.2%)	24 (43.6%)	178 (45.2%)	125 (26.8%)
Tumor stage						
T1–2	28	108 (62.1%)	63 (51.2%)	45 (81.8%)	244 (61.9%)	322 (69.1%)
T3–4	14	66 (37.9%)	60 (48.8%)	10 (18.2%)	150 (38.1%)	144 (30.9%)
Clinical N						
N0	10	59 (33.9%)	46 (37.4%)	19 (34.5%)	134 (34.0%)	155 (33.3%)
N1–3	16	115 (66.1%)	77 (62.6%)	36 (65.5%)	244 (61.9%)	311 (66.7%)
Unknown	16	0 (0%)	0 (0%)	0 (0%)	16 (4.1%)	0 (0%)
Histologic						
G1–2	18	118 (67.8%)	63 (51.2%)	52 (94.5%)	251 (63.7%)	401 (86.1%)
G3	16	45 (25.9%)	54 (43.9%)	3 (5.5%)	118 (29.9%)	63 (13.5%)
Unknown	8 (19.0%)	11 (6.3%)	6 (4.9%)	0 (0%)	25 (6.4%)	2 (0.4%)
ER score	—	—	—	—	—	
3–6	—	—				62 (13.3%)
7–8						404 (86.7%)
PR status	—					
Negative		46 (26.4%)	30 (24.4%)	17 (30.9%)	93 (23.6%)	82 (17.6%)
Positive		128 (73.6%)	93 (75.6%)	38 (69.1%)	259 (65.7%)	384 (82.4%)
Unknown		0 (0%)	0 (0%)	0 (0%)	42 (10.7%)	0 (0%)
Ki-67	—	—	—	—	—	
<20%						133 (28.6%)
≥20%						290 (62.2%)
Unknown						43 (9.2%)
NCT response						
pCR	10	11 (6.3%)	19 (15.5%)	5 (9.1%)	45 (11.4%)	22 (4.7%)
No pCR	32	162 (93.1%)	94 (76.4%)	50 (90.9%)	338 (85.8%)	438 (94.0%)
Unknown	0 (0%)	1 (0.6%)	10 (8.1%)	0 (0%)	11 (2.8%)	[e]6 (1.3%)

Table 1. Clinical and pathological characteristics of the public and Asan cohorts

(Continued)

Characteristics	Public data series from GEO					Asan cohort
	MADCC IGR	MADCC I-SPY-1	MADCC LBJ, INEN GEICAM USO	Miyake	Total	
[f]DRFS data	—			—		
Censored		151 (86.8%)	105 (85.4%)		256 (86.2%)	382 (82.0%)
Yes		23 (13.2%)	18 (14.6%)		41 (13.8%)	84 (18.0%)
[g]OS data	—	—	—	—	—	
Censored						414 (88.8%)
Yes						52 (11.2%)
DNA	U133A	U133A	U133A	U133P2		NA

[a]GSE22093, anthracycline-based NCT. [b]GSE25055, taxane-anthracycline NCT. [c]GSE25065, taxane-anthracycline NCT. [d]GSE32646, taxane-anthracycline NCT. Empty cells containing an em-dash (—) represent no available data. [e]Number of patients who did not undergo surgical treatment because of loss to follow-up (n = 4), refusal of surgery (n = 1), and distant relapse (n = 1). [f]Number of patients with distant metastasis or death. [g]Number of patients with death. AC = anthracycline and cyclophosphamide; ACT = anthracycline, cyclophosphamide, and taxane; AT = anthracycline and taxane; FAC = 5-fluorouracil, doxorubicin, and cyclophosphamide; FEC = 5-fluorouracil, epirubicin, and cyclophosphamide; F/U = follow-up; GEO = Gene Expression Omnibus; PR,

progesterone receptor = TFAC = paclitaxel followed by 5-fluorouracil, doxorubicin, and cyclophosphamide; TFEC = paclitaxel followed by 5-fluorouracil, epirubicin, and cyclophosphamide.

Table 2. Association between ^{18}F-fluorodeoxyglucose metabolism of breast cancer and other characteristics[a]

Characteristics	Public cohorts (n = 394)		Asan cohort (n = 466)	
	^{18}F-FDG signature score	***P* value**	**Maximum SUV**	***P* value**
Age, years				
20–50 vs. >50	0.02 (-0.03–0.06)	0.52	0.09 (-0.59–0.77)	0.80
Tumor stage				
T1–2 vs. T3–4	0.02 (-0.03–0.07)	0.35	0.36 (-0.29–1.01)	0.28
Clinical N stage				
N0 vs. N1–3	0.04 (-0.01–0.09)	0.14	1.03 (0.40–1.66)	0.001
Histologic grade				
G1–2 vs. G3	0.20 (0.15–0.25)	< 0.001	3.19 (2.35–4.02)	< 0.001
ER score (Allred)				
3–6 vs. 7–8	—		-2.87 (-3.71–(-2.02))	< 0.001
PR status				
Negative vs. Positive	-0.12 (-0.18–(-0.07))	< 0.001	-1.55 (-2.32–(-0.77))	< 0.001
Ki-67 expression				
<20% vs. ≥20%	—		1.63 (0.99–2.27)	< 0.001

[a]Data represent the regression coefficients (95% confidence interval).

Em-dash (—), not available

Abbreviations: ER, estrogen receptor; ^{18}F-FDG, ^{18}F-fluorodeoxyglucose; PR, progesterone receptor.

Table 3. Comparison of clinical and pathological characteristics between patients who did and did not undergo ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography before neoadjuvant chemotherapy

Characteristics	^{18}F-FDG PET/CT (n = 466)	No ^{18}F-FDG PET/CT (n = 55)	*P* value
Age, years			
20–50	341 (73.2%)	45 (81.8%)	0.22
>50	125 (26.8%)	10 (18.2%)	
Tumor stage			
T1–2	322 (69.1%)	38 (69.1%)	1.00
T3–4	144 (30.9%)	17 (30.9%)	
Clinical N stage			
N0	155 (33.3%)	17 (30.9%)	0.84
N1–3	311 (66.7%)	38 (69.1%)	
Histologic grade			0.72
G1–2	401 (86.4%)	46 (83.6%)	
G3	63 (13.6%)	8 (14.6%)	
Unknown	2 (0%)	1 (1.8%)	
ER score			1.00
3–6	62 (13.3%)	7 (12.7%)	
7–8	404 (86.7%)	48 (87.3%)	
PR status			0.08
Negative	82 (17.6%)	4 (7.3%)	
Positive	384 (82.4%)	51 (92.7%)	
Ki-67 expression			
<20%	133 (28.5%)	17 (30.9%)	0.31
≥20%	290 (62.2%)	24 (43.6%)	
Unknown	43 (9.2%)	14 (25.5%)	

Abbreviations: ER, estrogen receptor; ^{18}F-FDG, ^{18}F-fluorodeoxygluose; PET/CT, positron emission tomography/computed tomography; PR, progesterone receptor.

Table 4. Univariable analysis for pathological complete response of the public cohorts (n = 383)

Characteristics	Odds ratio (95% CI)	*P* value
Age, years: 20–50 vs. >50	2.00 (0.64–2.24)	0.57
Tumor stage: T1–2 vs. T3–4	1.49 (0.79–2.79)	0.21
Clinical N stage: N0 vs. N1–3	1.35 (0.66–2.77)	0.41
Histologic grade: G1–2 vs. G3	5.12 (2.58–10.17)	< 0.001
PR status: negative vs. positive	0.57 (0.28–1.15)	0.12
^{18}F-FDG signature score	8.38 (2.58–27.38)	< 0.001
21-gene recurrence score	14.61 (4.85–45.64)	< 0.001
70-gene signature score	32.64 (9.41–120.93)	< 0.001
50-gene score	17.48 (4.54–77.60)	< 0.001
12-gene risk score	28.92 (7.57–142.33)	< 0.001
76-gene signature score	3.45 (1.11–10.49)	0.03
Genomic grade index	19.70 (6.25–66.48)	< 0.001
CIN70 signature score	19.80 (6.01–69.42)	< 0.001
PTEN signature score	20.99 (6.40–74.02)	< 0.001
Proliferation gene module	9.65 (2.85–34.40)	< 0.001
ESR1 signature score	0.11 (0.03–0.33)	< 0.001

Abbreviations: CI, confidence interval; CIN70, chromosomal instability 70 genes; ESR, estrogen receptor expression; ^{18}F-FDG; 18F-fluorodeoxyglucose; NA, not applicable; PR, progesterone receptor; PTEN, phosphatase and tensin homolog.

Table 5. Multivariable analysis for pathological complete response of the public cohorts (n = 383)

Models	Odds ratio (95% CI), *P* value	
	[a]Multigene signature	[b]HG (G1–2 vs. G3)
^{18}F-FDG signature score + HG	3.12 (0.85–11.45), *P* = 0.09	4.09 (1.96–8.57), *P* < 0.001
21-gene recurrence score + HG	6.51 (1.94–21.85), *P* = 0.003	3.45 (1.64–7.25), *P* = 0.001
70-gene signature score + HG	14.17 (3.48–57.76), *P* < 0.001	2.90 (1.36–6.20), *P* = 0.006
50-gene score + HG	6.26 (1.36–28.83), *P* = 0.02	3.51 (1.66–7.40), *P* = 0.001
12-gene risk score + HG	12.71 (2.80–57.68), *P* = 0.001	2.90 (1.38–6.10), *P* = 0.005
76-gene signature score + HG	1.53 (0.44–5.32), *P* = 0.51	4.78 (2.33–9.81), *P* < 0.001
Genomic grade index + HG	8.86 (2.39–32.83), *P* = 0.001	3.04 (1.43–6.48), *P* = 0.004
CIN70 signature score + HG	8.32 (2.15–32.11), *P* = 0.002	3.19 (1.50–6.79), *P* = 0.003
PTEN signature score + HG	9.01 (2.35–34.53), *P* = 0.001	3.10 (1.46–6.59), *P* = 0.003
Proliferation gene module + HG	3.40 (0.83–13.90), *P* = 0.09	3.93 (1.85–8.32), *P* < 0.001
ESR1 signature score + HG	0.22 (0.07–0.71), *P* = 0.01	4.15 (2.03–8.49), *P* < 0.001

[a]Odds ratio of multigene signature included in each multivariable model.

[b]Odds ratio of histologic grade included in each multivariable model.

Abbreviations: CI, confidence interval; CIN70, chromosomal instability 70 genes; ESR, estrogen receptor expression; ^{18}F-FDG; 18F-fluorodeoxyglucose; HG, histologic grade; PTEN, phosphatase and tensin homolog.

Table 6. Univariable analysis for pathological complete response of the Asan cohort (n = 460)

Characteristics		Odds ratio (95% CI)	*P* value
Age, years: 20–50 vs. >50		0.81 (0.26–2.09)	0.68
Tumor stage: T2 vs. T3–4		1.06 (0.40–2.57)	0.90
Clinical N stage: N0 vs. N1–3		1.35 (0.54–3.82)	0.54
Histologic grade: G1–2 vs. G3		2.58 (0.97–6.89)	0.06
ER score (Allred): 3–6 vs. 7–8		0.50 (0.19–1.56)	0.19
PR status: negative vs. positive		0.19 (0.08–0.45)	< 0.001
Ki-67 expression: <20% vs. ≥20%		1.73 (0.56–5.38)	0.34
Maximum SUV, continuous		1.01 (0.87–1.13)	0.93
[a]Maximum SUV, <5.14 vs. ≥ 5.14		0.84 (0.35–1.99)	0.69
Maximum SUV,	Ter1 vs. Ter2	0.66 (0.22–1.87)	0.44
	Ter1 vs. Ter3	0.77 (0.27–2.13)	0.62

[a]Maximum SUV was dichotomized by the median value.

Abbreviations: CI, confidence interval; DRFS, distant relapse-free survival; PR, progesterone receptor; SUV, standardized uptake value of ^{18}F-fluorodeoxyglucose; Ter1, low tertile of SUV (1.36–4.14); Ter2, middle tertile of SUV (4.14–6.62); Ter3, high tertile of SUV (6.70–25.06).

Table 7. Univariable Cox proportional hazard analysis for distant relapse-free survival of the public cohort (n = 297)

Characteristics	Hazard ratio (95% CI)	C-index (95% CI)	*P* value
Age, years: 20–50 vs. >50	0.60 (0.32–1.14)	0.55 (0.47–0.63)	0.12
Tumor stage, T1–2 vs. T3–4	1.39 (0.75–2.56)	0.57 (0.49–0.65)	0.29
Clinical N stage: N0 vs. N1–3	4.42 (1.73–11.26)	0.63 (0.57–0.69)	0.002
Histologic grade: G1–2 vs. G3	1.22 (0.65–2.30)	0.51 (0.43–0.59)	0.54
PR status: negative vs. positive	0.94 (0.47–1.87)	0.49 (0.42–0.56)	0.85
^{18}F-FDG signature score	5.02 (1.63–15.42)	0.64 (0.55–0.72)	0.005
21-gene recurrence score	5.26 (1.92–14.44)	0.67 (0.59–0.75)	0.001
70-gene signature score	5.50 (1.75–17.31)	0.62 (0.53–0.72)	0.004
50-gene score	11.36 (3.12–41.41)	0.68 (0.59–0.77)	< 0.001
12-gene risk score	4.27 (1.51–12.04)	0.64 (0.56–0.72)	0.006
76-gene signature score	1.77 (0.64–4.88)	0.54 (0.45–0.63)	0.27
Genomic grade index	3.12 (1.15–8.46)	0.61 (0.52–0.70)	0.03
CIN70 signature score	2.45 (0.83–7.23)	0.58 (0.49–0.67)	0.10
PTEN signature score	2.88 (1.01–8.19)	0.59 (0.50–0.69)	0.047
Proliferation gene module	2.13 (0.69–6.56)	0.55 (0.46–0.65)	0.19
ESR1 signature score	0.20 (0.07–0.58)	0.65 (0.56–0.74)	0.003

Abbreviations: CI, confidence interval; CIN70, chromosomal instability 70 genes; ESR, estrogen receptor expression; ^{18}F-FDG, ^{18}F-fluorodeoxyglucose; PR, progesterone receptor; PTEN, phosphatase and tensin homolog.

Table 8. Multivariable Cox proportional hazard analysis for distant relapse-free survival of the public cohorts (n = 297)

[a]Models	Hazard ratio (95% CI), *P* value	
	[b]Multigene signature	[c]Clinical N stage
^{18}F-FDG signature score + N stage	4.85 (1.57–15.05), *P* = 0.006	4.33 (1.70–11.05), *P* = 0.002
21-gene recurrence score + N stage	5.75 (2.01–16.46), *P* = 0.001	4.48 (1.75–11.45), *P* = 0.002
70-gene signature score + N stage	6.45 (1.93–21.55), *P* = 0.002	4.60 (1.80–11.77), *P* = 0.001
50-gene score + N stage	10.49 (2.69–40.82), *P* = 0.001	3.97 (1.56–10.15), *P* = 0.004
12-gene risk score + N stage	3.68 (1.31–10.34), *P* = 0.01	4.10 (1.61–10.48), *P* = 0.003
76-gene signature score + N stage	1.71 (0.59–4.95), *P* = 0.32	4.38 (1.72–11.17), *P* = 0.002
Genomic grade index + N stage	3.37 (1.15–9.85), *P* = 0.03	4.44 (1.74–11.33), *P* = 0.002
CIN70 signature score + N stage	2.61 (0.84–8.08), *P* = 0.10	4.45 (1.75–11.36), *P* = 0.002
PTEN signature score + N stage	2.84 (0.96–8.41), *P* = 0.06	4.35 (1.71–11.09), *P* = 0.002
Proliferation gene module + N stage	2.10 (0.66–6.69), *P* = 0.21	4.39 (1.72–11.20), *P* = 0.002
ESR1 signature score + N stage	0.21 (0.07–0.63), *P* = 0.005	4.27 (1.67–10.88), *P* = 0.002

[a]Models including multigene signature and clinical N stage.

[b]Hazard ratio of multigene signature included in each multivariable model.

[c]Hazard ratio of clinical N stage (N0 vs. N1–3) in each multivariable model.

Abbreviations: CI, confidence interval; CIN70, chromosomal instability 70 genes; ESR, estrogen receptor expression; ^{18}F-FDG, ^{18}F-fluorodeoxyglucose; PTEN, phosphatase and tensin homolog.

Table 9. C-index analyses of prognostic models for distant relapse-free survival of the public cohorts (n = 297)

Models	**C-index (95% CI)**	[a]***P* value**
Clinical N stage	0.63 (0.57–0.69)	NA
^{18}F-FDG signature score + clinical N stage	0.70 (0.62–0.78)	0.009
21-gene recurrence score + clinical N stage	0.72 (0.65–0.80)	< 0.001
70-gene signature score + clinical N stage	0.70 (0.62–0.78)	0.03
50-gene score + clinical N stage	0.73 (0.65–0.81)	0.001
12-gene risk score + clinical N stage	0.70 (0.62–0.78)	< 0.001
76-gene signature score + clinical N stage	0.64 (0.56–0.72)	0.63
Genomic grade index + clinical N stage	0.68 (0.60–0.77)	0.04
CIN70 signature score + clinical N stage	0.67 (0.59–0.76)	0.11
PTEN signature score + clinical N stage	0.68 (0.60–0.76)	0.09
Proliferation gene module + clinical N stage	0.65 (0.57–0.73)	0.42
ESR1 signature score + clinical N stage	0.71 (0.63–0.80)	0.003
^{18}F-FDG + ESR1 signature score + clinical N stage	0.72 (0.64–0.80)	0.003

[a]*P* values indicate improved prediction relative to clinical N stage alone. CIN70 = chromosomal instability 70 genes; ESR1 = estrogen receptor 1; ^{18}F-FDG = ^{18}F-fluorodeoxyglucose; PTEN = phosphatase and tensin homolog.

Table 10. Univariable Cox proportional hazard analyses for survival of the Asan cohort (n = 466)

Characteristics	DRFS		OS	
	HR (95% CI)	***P* value**	**HR (95% CI)**	***P* value**
Age, years: 20–50 vs. >50	1.64 (1.05–2.57)	0.03	2.12 (1.22–3.70)	0.008
Tumor stage: T2 vs. T3–4	1.74 (1.13–2.68)	0.01	1.94 (1.12–3.36)	0.02
Clinical N stage: N0 vs. N1–3	4.03 (2.08–7.81)	< 0.001	8.56 (2.67–27.47)	< 0.001
Histologic grade: G1–2 vs. G3	1.07 (0.58–1.98)	0.82	1.73 (0.89–3.38)	0.11
ER score (Allred): 3–6 vs. 7–8	0.90 (0.49–1.65)	0.73	0.75 (0.36–1.53)	0.42
PR status: negative vs. positive	0.64 (0.39–1.06)	0.09	0.50 (0.28–0.91)	0.02
Ki-67 expression: <20% vs. ≥20%	2.48 (1.26–4.88)	0.009	4.63 (1.53–13.98)	0.007
Maximum SUV, continuous	1.08 (1.02–1.14)	0.01	1.13 (1.06–1.21)	< 0.001
[a]Maximum SUV, <5.14 vs. ≥ 5.14	2.21 (1.40–3.48)	0.001	3.70 (1.94–7.07)	< 0.001
Maximum SUV, Ter1 vs. Ter2	2.42 (1.26–4.63)	0.008	2.78 (1.00–7.71)	0.05
Maximum SUV, Ter1 vs. Ter3	3.48 (1.87–6.50)	< 0.001	7.19 (2.80–18.45)	< 0.001

[a]Maximum SUV was dichotomized by the median value.

Abbreviations: CI, confidence interval; DRFS, distant relapse-free survival; HR, hazard ratio; OS, overall survival; PR, progesterone receptor; SUV, standardized uptake value of ^{18}F-fluorodeoxyglucose; Ter1, low tertile of SUV (1.36–4.14); Ter2, middle tertile of SUV (4.14–6.62); Ter3, high tertile of SUV (6.70–25.06).

Table 11. Multivariable Cox proportional hazard analyses for distant relapse-free and overall survival of the Asan cohort (n = 466)

Characteristics		DRFS		OS	
		Hazard ratio (95% CI)	***P* value**	**Hazard ratio (95% CI)**	***P* value**
Age, years: 20–50 vs. >50		1.50 (0.95–2.37)	0.09	2.01 (1.13–3.58)	0.02
Tumor stage: T2 vs. T3-4		1.66 (1.06–2.58)	0.03	1.81 (1.02–3.20)	0.04
Clinical N stage: N0 vs. N1-3		3.01 (1.54–5.89)	0.001	6.21 (1.91–20.18)	0.002
Histologic grade: G1-2 vs. G3		0.76 (0.41–1.43)	0.40	1.06 (0.53–2.11)	0.87
PR status: negative vs. positive		0.75 (0.45–1.25)	0.27	0.58 (0.31–1.07)	0.08
Ki-67 expression: <20% vs. ≥20%		1.75 (0.88–3.48)	0.11	3.11 (0.97–9.99)	0.06
[ab]Maximum SUV	Ter1 vs. Ter2	2.26 (1.17–4.39)	0.02	3.01 (1.05–8.58)	0.04
	Ter1 vs. Ter3	2.93 (1.55–5.54)	0.001	6.62 (2.49–17.58)	< 0.001

[a]Hazard ratio (95% CI) of maximum SUV as a continuous measurement: 1.01 (0.99 to 1.12, $P = 0.09$) for distant relapse-free survival, and 1.12 (1.04–1.20, $P = 0.003$) for overall survival. [b]Hazard ratio (95% CI) of maximum SUV as a categorical estimate according to the median value (<5.14 vs. ≥5.14): 2.00 (1.25–3.18, $P = 0.004$) for distant relapse-free survival and 3.58 (1.82–7.05, $P < 0.001$) for overall survival. CI = confidence interval; PR = progesterone receptor; SUV = standardized uptake value of ^{18}F-fluorodeoxyglucose; Ter1 = low tertile of SUV (1.36–4.14); Ter2 = middle tertile of SUV (4.14–6.62); Ter3 = high tertile of SUV (6.70–25.06).

국문요약

목적: 에스트로겐 수용체 양성, ERBB2 음성 유방암 여성에서 ^{18}F-fluorodeoxyglucose (^{18}F-FDG) 대사와 원거리 무 재발 생존율 (DRFS)과의 연관성을 알아보고자 하였다.

재료 및 방법: 이 연구는 유방암 환자들의 공개 및 임상 코호트 자료들을 이용한 다중 코호트 연구이다. 2019 년 5 월까지의 microarray expression 이 포함된 공개 코호트 자료들은 미국과 일본에 위치한 3 개의 센터에서 수집되었다. 2007 년 11 월부터 2014 년 12 월까지의 서울 아산 병원 유방암 코호트 자료들 (아산 코호트)도 이 연구를 위해 수집되었다. 공개 및 아산 코호트 모두 anthracycline 기반의 수술 전 보조 화학 요법을 받은 유방암 여성 환자들이 이 연구에 포함되었다. 이 연구의 일차적인 목적은 DRFS 이다. 공개 코호트 자료에서 유방암에 대한 12 개의 다중 유전자 점수 (^{18}F-FDG signature 점수 포함)를 계산하였고 아산 코호트 자료의 ^{18}F-FDG 양전자 방출 단층 촬영 (PET)에서 유방암 병변들의 maximum standardized uptake value (maximum SUV)를 측정하였다. 추

가적으로, 종양의 ^{18}F-FDG 대사와 ^{18}F-FDG signature 점수 사이의 상관 관계의 분석을 위해 ^{18}F-FDG PET 과 유전자 발현 프로파일을 모두 가진 폐암환자가 포함된 별도의 공개 코호트를 이 연구에 포함시켰다.

결과: 분석에는 공개 코호트에서 394 명 (이 중 생존 데이터가 있는 여성은 297 명)과 아산 코호트에서 466 명의 여성들이 포함되었다. 추적 기간의 중앙값 (사분위수)는 각각 공개 코호트에서 3.2 (2.3-4.4) 및 아산 코호트에서 6.2 (5.3-7.6) 년이었다. 공개 코호트 자료에서 ^{18}F-FDG signature 점수는 DRFS 예측의 의미 있는 독립 변수였다 (HR, 4.85; 95 % CI, 1.57-11.05; P = 0.006). 그리고 예후 예측 성능도 (C-index: 0.70; 95 % CI, 0.62- 0.78) 다른 다중 유전자 점수와 비슷한 값을 보였다. 아산 코호트의 삼분위수로 나눈 종양 ^{18}F-FDG 대사 (maximum SUV)에 대한 다 변량 분석에서 (HR [95 % CI]) maximum SUV 의 중간 및 위 삼분위수들이 DRFS (2.26 [1.17-4.39], P = 0.02; 및 2.93 [1.55-5.54], P = 0.001)의 의미 있는 예후 인자였다. 8 년 DRFS 의 경우 낮은 삼분위수에서는 90.7 % (95 % CI, 85.5-96.1 %)로 높은 생존을 보인 반면 중간 및 높은 삼분위

수에서는 73.7 % (95 % CI, 68.0-79.8 %)로 상대적으로 낮은 생존을 보였다.

결론: ^{18}F-FDG PET는 ER 양성, ERBB2 음성 환자에서 원격 전이 및 사망 위험을 평가할 수 있다.

중심단어: 유방암, 에스트로겐 수용체, 포도당 대사, 생존, 양전자 방출 단층 촬영