

# A multiomics signature highlights alterations underlying homologous recombination deficiency in triple-negative breast cancer

Chao You (✉ [youchao8888@aliyun.com](mailto:youchao8888@aliyun.com))

Fudan University Shanghai Cancer Center, Fudan University

**Guan-Hua Su**

Fudan University Shanghai Cancer Center, Fudan University

**Lin Jiang**

Fudan University Shanghai Cancer Center, Fudan University

**Yi Xiao**

Fudan University Shanghai Cancer Center, Fudan University

**Ren-Cheng Zheng**

Fudan University

**He Wang**

Fudan University

**Yi-zhou Jiang**

Fudan University Shanghai Cancer Center, Fudan University

**Wei-jun Peng**

Fudan University Shanghai Cancer Center, Fudan University

**Zhi-Ming Shao**

Fudan University Shanghai Cancer Center, Fudan University

**Ya-Jia Gu**

Fudan University Shanghai Cancer Center, Fudan University

---

## Research Article

**Keywords:** homologous recombination deficiency, triple-negative breast cancer, multiomics, radiomics

**Posted Date:** January 18th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1244930/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** Homologous recombination (HR) is a key pathway in DNA double-strand damage repair. HR deficiency (HRD) more commonly occurs in triple-negative breast cancers (TNBCs) than in other breast cancer subtypes. Several clinical trials have demonstrated the value of HRD in stratifying breast cancer patients into distinct response groups to poly ADP ribose polymerase (PARP) inhibitors and chemotherapy.

**Methods:** We retrospectively collected TNBC samples to establish a multiomics cohort (n=343) to explore the biological and phenotypic mechanisms underlying the better prognosis of patients with high HRD scores. The HRD score was calculated using the sum of the telomeric allelic imbalance (TAI) score, loss of heterozygosity (LOH) score and large-scale state transition (LST) score. Gene set enrichment analysis (GSEA) were conducted to reveal underlying pathways of low HRD score patients. Radiomics model were established to predict HRD score in a noninvasive method.

**Results:** Multivariable Cox analysis revealed the independent prognostic value of a low HRD score (hazard ratio, 2.20, 95% confidence interval, 1.05 to 4.59, p=0.04). Furthermore, amino acid and lipid metabolism pathways were highly enriched in tumors with low HRD scores, which was also demonstrated by differentially abundant metabolite analysis. A noninvasive radiomics method was developed to predict HRD status and performed well in the independent validation cohort (support vector machine model: area under the curve (AUC), 0.739, sensitivity, 0.571, and specificity, 0.824; logistic regression model: AUC, 0.695, sensitivity, 0.571, and specificity, 0.882).

**Conclusions:** In conclusion, we revealed the prognostic value of the HRD score, predicted HRD status with noninvasive radiomics features and preliminarily explored druggable targets in TNBC patients with low HRD scores.

## Background

Triple-negative breast cancer (TNBC) is a subtype of breast cancer defined by the absence of estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2)<sup>1,2</sup>. TNBC typically occurs in young patients and progresses rapidly, leading to poor prognosis<sup>3,4</sup>. The lack of effective therapeutic targets indicates an urgent need to identify novel targets and screen patients who are suitable for specific therapies. In recent years, the development of sequencing techniques has enabled us to unveil the heterogeneity of breast cancer from the distinct biological characteristics of several subtypes<sup>5-7</sup>. Our previous study classified TNBC patients into four subtypes, luminal androgen receptor (LAR), basal-like immune-suppressed (BLIS), immunomodulatory (IM), and mesenchymal-like (MES), which offer practicable treatment strategies<sup>5-7</sup>.

DNA in cellular nuclei is continuously exposed to risk factors for damage, causing a variety of patterns of lesions, including single-strand breaks (SSBs), double-strand breaks (DSBs), base mismatches, insertions

and deletions<sup>8</sup>. Among them, DSBs are typically repaired by homologous recombination (HR) and nonhomologous end joining (NHEJ)<sup>9</sup>. Cancer cells harboring mutations in one of the genes orchestrating HR, such as the BRCA1, BRCA2, RAD51 and PALB2 genes, are more vulnerable to DNA-damaging drug platinum agents and SSB repair suppressor poly ADP ribose polymerase (PARP) inhibitors<sup>10-13</sup>. Notably, Timms *et al.* introduced the HR deficiency (HRD) score in a retrospective study by summing telomeric allelic imbalance (TAI), loss of heterozygosity (LOH) and large-scale state transition (LST)<sup>14-17</sup>. Telli *et al.* defined HRD by HRD score  $\geq 42$  or BRCA 1/2 mutation and found that breast cancer patients with HRD were prone to obtaining pathological complete response (pCR) to neoadjuvant chemotherapy. In addition, a high HRD score (HRD  $\geq 42$ ) but BRCA 1/2 wildtype was positively correlated with a higher pCR rate, further demonstrating the value of the HRD score in breast cancer<sup>18</sup>. Data from The Cancer Genome Atlas (TCGA) revealed that germline BRCA1/2 mutations exist in approximately 10% of patients with TNBC<sup>19-21</sup>. However, the mechanism underlying the better clinical outcomes of high-HRD TNBC patients has not been thoroughly investigated.

In this study, we integrated multiomics information from our TNBC cohort at Fudan University Shanghai Cancer Center (FUSCC). Genomics, transcriptomics, metabolomics and radiomics were utilized to calculate the HRD score, analyze the difference between high- and low-HRD score patients and noninvasively predict HRD status.

## Methods

### Study Population

A total of 343 TNBC patients with preoperative MRI, primary tumor tissue and blood samples obtained from FUSCC made up the final study group. All patients underwent surgery at FUSCC between 2007 and 2014 with a median follow-up duration of 47.3 months. A total of 67 patients experienced relapse and/or metastatic progression during this period. The clinical data included age, tumor size, lymph node status and Ki-67 index.

In this study, all 343 patients were included in the HRD analysis. Among them, 118 patients were available for radiomics analysis and had preoperative breast contrast-enhanced (CE)-MRI, and 252 samples of hematoxylin and eosin (H.E.)-stained sections with immunohistochemistry (IHC) staining data were available. In addition, 252 samples of transcriptomic sequencing data and 246 samples of polar metabolomics and lipidomics data were also eligible for analysis. The generation of transcriptomics and metabolomics data has been described previously<sup>22,23</sup>. All the samples from TNBC patients were used for (1) HRD analysis and prognosis assessment (n=343); (2) radiomics signature analysis and validation for HRD identification (n=118); and (3) molecular and biological assessment of associations between HRD and other multiomics, including pathology, transcriptomics and metabolomics (n=252, 252 and 246, respectively) (**Supplementary Figure 1A**). This study was approved by the Medical Ethics Committee of FUSCC, and written informed consent was provided by each patient. The study was conducted in accordance with the Declaration of Helsinki as revised in 2013.

## HRD Analysis

As summarized in a previous study <sup>14</sup>, the HRD score was calculated as the arithmetic sum of three scores using segments with integer copy numbers produced by ASCAT: TAI score, LOH score and LST score. The calculation of these three scores was described in a previous article <sup>14</sup>. In brief, the TAI score was defined as the number of regions with allelic imbalance that were longer than 11 Mb and extended to one of the subtelomeres but did not cross the centromere. The LOH score was defined as the number of LOH regions longer than 15 Mb but shorter than the whole chromosome. The LST score was defined as the number of break points between regions longer than 10 Mb after filtering out regions shorter than 3 Mb. To diminish the effect of ploidy, the LST score was modified according to the formula  $LST_m = LST - kP$ , where P is ploidy and k is a constant of 15.5. The HRD cohort was divided into a high-HRD (HRD score  $\geq 42$ ) cohort and a low-HRD cohort (HRD score  $< 42$ ) <sup>18</sup>.

## Analysis of Differentially Abundant Metabolites and Differentially Expressed Genes

The differentially expressed genes were determined through the limma R package <sup>24</sup>. The differential expression analysis outputs of limma were used to generate the ranked list file. GSEA was performed using the clusterProfiler R package <sup>25</sup>. The differential abundance of metabolites was calculated by performing the Mann-Whitney U tests for all the detected metabolites.

## MRI Technique and Radiomics Analysis

MRI scans were obtained using a 1.5-T dedicated spiral breast MRI system (Aurora Imaging Technology, Aurora Systems, Inc., Vancouver, Canada) described elsewhere. In this study, CE-MRI data were used for radiomic analysis. The T1-weighted fat suppressed sequence (time of echo/time of repetition, 4.8/29 ms; thickness, 1.1 mm; field of view, 360×360 mm; matrix, 360× 360) included one precontrast image and four postcontrast images following a bolus injection of gadopentetate dimeglumine at 2 ml/s (Magnevist, Bayer Schering Pharma, Berlin, Germany) using an automatic injector. Postcontrast images were obtained 90, 180, 270, and 360 seconds after injection.

The whole tumor was semiautomatically segmented on the peak enhanced phase of CE-MRI by 3D slicer software (<https://www.slicer.org/>). The tissue surrounding the tumor or inside the tumor was also identified because it can provide information on both the microenvironment and the tumor itself. Specifically, the peritumoral area was obtained by expanding the tumor outward by a 2-pixel width (0.7 mm × 2 pixels = 1.4 mm) and subtracting the tumor area, while the intratumoral area was obtained by shrinking by a 2-pixel width.

Spatial domain feature extraction, including shape features, first-order features, texture features was performed for each sample using the PyRadiomics package. Sequential features were calculated based on each spatial domain extracted from each contrasted-enhanced phase (Detailed description of radiomics features was presented in **Supplementary Materials**). Feature selection was performed using the least absolute shrinkage and selection operator (Lasso) method (glmnet R package). Tuning

parameter ( $\lambda$ ) selection in the Lasso model used 10-fold cross validation. The selected radiomics features were calculated for each patient by three different methods, including multivariate linear regression (glm R package), SVM (e1071 R package), and deep neural network (sklearn Python package). Radiomics prediction models were validated internally and externally. First, the trained classifiers were assessed by 10-fold cross validation via the glmnet R package<sup>26</sup>. Then, the trained classifiers were further tested in the validation cohort, and the AUC of the ROC curve was assessed.

## Statistical Analysis

Student's t test and Wilcoxon's test were applied to compare continuous variables. Prior to the comparisons, the normality of the distribution of variables was tested with the Shapiro-Wilk test. Pearson's chi-square test and Fisher's exact test were employed for the comparison of unordered categorical variables. To explore the association between the HRD score and survival, Kaplan–Meier analysis and a Cox proportional hazards model were employed. Comparison of survival between groups was conducted via the log rank test. The predictive radiomics HRD model was validated by using the SVM and LR classifiers. First, the trained classifiers were assessed by 10-fold cross validation to reduce potential overfitting. Then, the trained classifiers were further tested in the validation cohort, and the AUC of the ROC curve was assessed. The ROC value was determined by the Youden Index. In multiple hypothesis testing, false discovery rate (FDR) correction was used to decrease false positive rates. All the tests were two-sided, and  $P < 0.05$  was regarded as indicating significance unless otherwise stated. All of the statistical analyses were performed with R software (version 3.6.1, <http://www.R-project.org>).

## Results

### Study Population and Clinical Characteristics

A total of 343 TNBC patients with HRD scores were included in the present study. Of these, 135 patients (39.4%) had a high HRD score, and 208 (60.6%) had a low HRD score. Age and Ki-67 index were found to differ significantly between the high-HRD and low-HRD groups ( $p=0.02$  and  $0.0004$ , respectively). No significant differences in tumor size or lymph node status were observed. The clinicopathological characteristics of the TNBC patients are listed in **Table 1**.

### Prognostic Value of HRD in TNBC

Univariate Cox proportional hazards analysis showed that patients with high HRD scores had a higher probability of recurrence-free survival (RFS) in the overall TNBC cohort ( $p=0.03$ ) (**Figure 1A**). Among the various TNBC subtypes, high HRD showed a better RFS rate in only the BLIS subtype ( $p=0.002$ ) but not in the other subtypes (**Figure 1B**). No significant correlations were observed between HRD and prognosis in the basal and non-basal subtypes (**Figure 1C-D**). A multivariate Cox proportional hazards model also revealed that HRD together with T and N status independently predicted RFS in all TNBC patients. Stage T3/T4, stage N2/N3 and low HRD were positively associated with worse RFS (**Figure 1E**).

Correlation analysis indicated that cluster 3, characterizing immune-inflamed TNBC revealed by our previous study, was more enriched in high-HRD patients (**Figure 2A**)<sup>27</sup>. Notably, for the distinct subtypes of TNBC, the LAR subtype showed lower HRD scores, while the BLIS subtype presented the opposite finding; basal-like breast cancers were more enriched in high-HRD group (**Figure 2A**)<sup>7</sup>. In addition, high-HRD patients exhibited a higher tumor mutation burden, higher lymphocyte infiltration, and fewer positive lymph nodes (**Figure 2B**).

### **Transcriptomic and Metabolomic Characteristics of TNBCs with Low HRD Scores**

Given the important role of the HRD score discussed in previous reports, we analyzed the transcriptomic differences between low- and high-HRD score patients with TNBC (**Figure 3A**). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis revealed that metabolic pathways were widely upregulated in low-HRD score patients (**Figure 3B**). Consistently, gene set enrichment analysis (GSEA) based on the KEGG and Reactome databases showed that metabolism-related pathways, particularly amino acid and lipid metabolism pathways, were significantly upregulated, while DNA damage repair- and cell cycle-associated pathways were downregulated in low-HRD patients (**Figure 3C-D**). Then, we conducted a metabolomic analysis of samples with different HRD scores. The metabolomics analysis indicated that amino acids, peptides and lipids were differentially expressed in patients with different HRD scores, which aligned with our findings from the transcriptomics analysis (**Figure 3E**). Furthermore, lipidomics data revealed that sphingolipids were the largest proportion of lipids upregulated in low-HRD score patients (**Figure 3F**).

### **Predicting HRD Based on Noninvasive Radiomics**

A total of 118 patients with preoperative MRI data included in the radiomic cohort were randomly assigned into a training cohort (n=80) and a validation cohort (n=38). Thirty-eight (32.2%) of these patients had high HRD scores, and 80 (67.8%) had low HRD scores.

To explore whether a noninvasive radiomics approach could identify HRD status with high accuracy, radiomic prediction models were curated via two steps: feature selection and model building. The HRD score was calculated as the arithmetic sum of three scores as the gold standard to evaluate the prediction efficacy<sup>14</sup>. Three sequential radiomics features were selected to predict the HRD status in the TNBC cohort (**Figure 4A**). They were the mean value of the first-order Minimum features after wavelet-LLL filtering from the tumor area, the kurtosis of the texture GLSZM Large Area High Gray Level Emphasis feature after wavelet-LHH filtering from the peritumoral region, and the difference of first post-enhanced phase and plain scan value of wavelet value of the first-order Maximum features after wavelet-HHH filtering from the peritumoral region. The area under the curve (AUC), sensitivity, and specificity of support vector machine (SVM) were 0.863, 0.797, and 0.857 in the training cohort, and these indices were 0.739, 0.571, and 0.824 in the validation cohort. The AUC, sensitivity, and specificity of logistic regression (LR) were 0.864, 0.831, and 0.810 in the training cohort, while the above indices were 0.695, 0.571, and 0.882 in the validation cohort (**Figure 5B**).

Age and Ki-67 were identified as variables for inclusion in the clinical model for predicting HRD status, which achieved AUCs of 0.65 and 0.63 in the training and validation cohorts, respectively. The radiomics model showed a significantly higher AUC than the clinical model ( $p < 0.05$ ), while the combined model showed no statistical significance in either cohort (**Supplementary Figure 1A-B**).

## Discussion

The HRD score is a useful index to evaluate HR and infer sensitivity to platinum chemotherapy and PARP inhibitors in TNBC patients<sup>14,18</sup>. Integrating the large-scale multiomics data of TNBC patients treated in our institution, we unraveled the possible mechanisms underlying the distinct prognosis between patients with different HRD scores and investigated the convenient clinical utilization of the HRD score in a noninvasive radiomics manner. Based on our multiomics database, we calculated the HRD score of each patient by arithmetically adding up the three scores reported previously<sup>14</sup>. Significantly better survival outcomes were observed in the high-HRD group of patients, and its role as an independent prognostic factor for TNBC was demonstrated by multivariable Cox regression. Transcriptomics analysis revealed that DNA damage repair- and cell cycle-associated pathways were upregulated in high-HRD patients, while low-HRD patients showed upregulation in more metabolic pathways, particularly amino acid and lipid metabolism. We therefore investigated the metabolites upregulated in low-HRD patients using matched metabolomics data. The results showed that amino acids, peptides and lipids, particularly sphingolipids, were highly expressed in low-HRD score patients, which provided a clue for targeting metabolites in low-HRD TNBC patients even though they could not benefit from platinum chemotherapy or PARP inhibitors. Given the invasive and time-consuming properties of HRD score detection in the clinic, we established a radiomics signature to invasively predict the HRD status of TNBC patients.

HR is one of the mechanisms of DNA DSB repair<sup>28</sup>. Tumors with higher HRD scores were considered to be more deficient in the process of HR, resulting in cell apoptosis. In the present study, we demonstrated that a higher HRD score correlated with longer RFS, which was consistent with the findings of previous clinical trials and studies. For instance, Sharma *et al.* conducted an analysis of TNBC patients treated with adjuvant doxorubicin and cyclophosphamide in the SWOG S9313 clinical trial and found that positive HRD (defined as HRD score  $\geq 42$  or tumor BRCA 1/2 mutation) was associated with longer disease-free survival (DFS) and overall survival (OS)<sup>29</sup>.

Radiomics is an emerging field for dissecting information from medical imaging through high-throughput radiomics features<sup>30-32</sup>. In this study, we adopted noninvasive radiomics methods to differentiate TNBC patients with high and low HRD scores, which will mitigate the cost and suffering from the invasive detection of HRD. Previous studies have revealed the application of radiomics in multiple aspects of the clinical context. Aerts *et al.* revealed the prognostic value and correlation with gene expression patterns of noninvasive radiomics features in lung cancer and head and neck cancer<sup>33</sup>. The efficiency of radiomics in predicting distant metastasis-free survival was tested in locally advanced rectal cancer<sup>34</sup>. A

24-feature radiomics signature was constructed to predict lymph node metastasis in colorectal cancer. A nomogram containing a radiomics signature performed well in lymph node prediction, with a C-index of over 0.75<sup>35</sup>. Notably, CD8-positive T cells and patient response to immunotherapy were predicted by radiomics features<sup>36,37</sup>. Our results together with those of previous radiomics studies provide support that radiomics has the potential to benefit clinical index prediction and decision making.

In a previous study, our research team focused on the multiomics characteristics of TNBC patients in an Asian population. Transcriptomics analysis partitioned TNBC patients into four subtypes: LAR, BLIS, IM and MES<sup>7</sup>. In the BLIS subtype, the HRD mutation signature was enriched, inferring that implementing platinum chemotherapy or PARP inhibitors in high-HRD patients and escalated therapy in low-HRD patients might be appropriate precision treatment strategies for BLIS TNBCs<sup>12,13,38-42</sup>. Based on the properties of each subtype, an umbrella trial was conducted, and an objective response was achieved in 29% of refractory metastatic TNBC patients<sup>43</sup>. However, acknowledged recommended agents are lacking in low-HRD TNBC patients. Furthermore, our transcriptomics and metabolomics analyses demonstrated metabolic heterogeneity in TNBC patients<sup>44</sup>. These studies indicated the significance of transcriptome and metabolome information in molecular portraits and heterogeneity in TNBCs. Consequently, we conducted multiomics analysis in patients with different HRD statuses. We verified that cell cycle and DNA damage repair pathways were enriched in high-HRD patients and found that amino acid and lipid metabolism pathways, particularly sphingolipid metabolism, were activated in low-HRD patients. These metabolic alterations might serve as druggable targets in low-HRD TNBC patients.

Several limitations were contained in our study. First, our analysis was based on a retrospective cohort of the largest multiomics dataset comprising transcriptomics, metabolomics, and radiomics data; however, the results need retrospective and prospective validation in external cohorts and clinical trials. Second, the possible mechanisms underlying the worse prognosis of low-HRD patients were first reported through a multiomics cohort with clinicopathological information. Experimental works are necessary to further validate this finding and investigate the explicit mechanisms.

## Conclusions

In conclusion, we identified a set of patients with higher HRD scores in the FUSCC multiomics TNBC cohort and demonstrated the prognostic value of the HRD score in TNBC. Notably, we proposed the practicability of predicting HRD status with a noninvasive radiomics model. Moreover, it was revealed that amino acid and lipid metabolism, particularly sphingolipid metabolism, was enriched in low-HRD tumors, providing potential therapeutic targets for low-HRD TNBC patients.

## Abbreviations

area under the curve (AUC)

basal-like immune-suppressed (BLIS)

contrast-enhanced (CE)

double-strand breaks (DSBs)

estrogen receptor (ER)

Fudan University Shanghai Cancer Center (FUSCC)

Gene set enrichment analysis (GSEA)

hematoxylin and eosin (H.E.)

Homologous recombination (HR)

HR deficiency (HRD)

human epidermal growth factor receptor 2 (HER2)

immunohistochemistry (IHC)

immunomodulatory (IM)

Kyoto Encyclopedia of Genes and Genomes (KEGG)

large-scale state transition (LST)

loss of heterozygosity (LOH)

luminal androgen receptor (LAR)

mesenchymal-like (MES)

nonhomologous end joining (NHEJ)

poly ADP ribose polymerase (PARP)

progesterone receptor (PR)

single-strand breaks (SSBs)

telomeric allelic imbalance (TAI)

The Cancer Genome Atlas (TCGA)

triple-negative breast cancers (TNBCs)

# Declarations

## Ethics approval and consent to participate

Institutional Review Board approval was obtained.

## Consent for publication

Written informed consent was obtained from all subjects (patients) in this study.

## Availability of data and materials

Raw LC-MS data have been uploaded to the MetaboLights dataset. The IDs were MTBLS3188 and MTBLS3189 for polar metabolomics and lipidomics, respectively. Microarray data and sequence data have been deposited in the NCBI Gene Expression Omnibus (OncoScan array; GEO: GSE118527) and Sequence Read Archive (WES and RNA-seq; SRA: SRP157974).

## Competing interests

The Authors declare no Competing Financial or Non-Financial Interests.

## Funding

This project was supported by the grants from The National Natural Science Foundation of China (81901703, 82071878), Cancer research program of national cancer center (Grant No. NCC201909B06), Youth Medical Talents-Clinical Imaging Practitioner Program [SHWRS(2020)\_087] and Clinical Research Plan of SHDC (SHDC2020CR2008A).

## Authors' contributions

All authors contributed fundamentally to this study and participated sufficiently to take public responsibility for its content. Chao You designed the study, Chao You and Guan-hua Su did the literature review and Guan-hua Su drafted the manuscript. Chao You, Guan-hua Su and Lin Jiang collected the data. Chao You and Guan-hua Su were responsible for quality control of data and interpretation. Lin Jiang and Guan-hua Su was involved in the statistical analysis and interpreted the results. Ren-cheng Zhen and He Wang was involved in the MR radiomics extraction. Chao You, Guan-hua Su, Yi Xiao and Yi-zhou Jiang interpreted the results and edited the manuscript. Wei-jun Peng and Zhi-ming Shao edited the manuscript. Chao You and Ya-jia Gu edited and finalized the manuscript. Publication is approved by all authors.

## Acknowledgements

We thank the staff of the Radiology Department of Fudan University Shanghai Cancer Center for their assistance in breast MRI imaging collection. Besides, we thank the staff of the Institute of Science and

Technology for Brain-inspired Intelligence of Fudan University for their contribution in radiomics features extraction.

## References

- 1 Waks, A. G. & Winer, E. P. Breast Cancer Treatment: A Review. *JAMA* **321**, 288-300, doi:10.1001/jama.2018.19323 (2019).
- 2 Criscitiello, C., Azim, H. A., Jr., Schouten, P. C., Linn, S. C. & Sotiriou, C. Understanding the biology of triple-negative breast cancer. *Ann Oncol* **23 Suppl 6**, vi13-18, doi:10.1093/annonc/mds188 (2012).
- 3 Denkert, C., Liedtke, C., Tutt, A. & von Minckwitz, G. Molecular alterations in triple-negative breast cancer—the road to new treatment strategies. *The Lancet* **389**, 2430-2442, doi:10.1016/s0140-6736(16)32454-0 (2017).
- 4 Metzger-Filho, O. *et al.* Dissecting the heterogeneity of triple-negative breast cancer. *J Clin Oncol* **30**, 1879-1887, doi:10.1200/JCO.2011.38.2010 (2012).
- 5 Lehmann, B. D. *et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest* **121**, 2750-2767, doi:10.1172/JCI45014 (2011).
- 6 Burstein, M. D. *et al.* Comprehensive genomic analysis identifies novel subtypes and targets of triple-negative breast cancer. *Clin Cancer Res* **21**, 1688-1698, doi:10.1158/1078-0432.CCR-14-0432 (2015).
- 7 Jiang, Y. Z. *et al.* Genomic and Transcriptomic Landscape of Triple-Negative Breast Cancers: Subtypes and Treatment Strategies. *Cancer Cell* **35**, 428-440 e425, doi:10.1016/j.ccell.2019.02.001 (2019).
- 8 Belli, C., Duso, B. A., Ferraro, E. & Curigliano, G. Homologous recombination deficiency in triple negative breast cancer. *Breast* **45**, 15-21, doi:10.1016/j.breast.2019.02.007 (2019).
- 9 Brown, J. S., O'Carrigan, B., Jackson, S. P. & Yap, T. A. Targeting DNA Repair in Cancer: Beyond PARP Inhibitors. *Cancer Discov* **7**, 20-37, doi:10.1158/2159-8290.CD-16-0860 (2017).
- 10 Yang, H. *et al.* BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. *Science* **297**, 1837-1848, doi:10.1126/science.297.5588.1837 (2002).
- 11 Pellegrini, L. *et al.* Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature* **420**, 287-293, doi:10.1038/nature01230 (2002).
- 12 Lord, C. J. & Ashworth, A. PARP inhibitors: Synthetic lethality in the clinic. *Science* **355**, 1152-1158, doi:10.1126/science.aam7344 (2017).

- 13 Ashworth, A. & Lord, C. J. Synthetic lethal therapies for cancer: what's next after PARP inhibitors? *Nat Rev Clin Oncol* **15**, 564-576, doi:10.1038/s41571-018-0055-6 (2018).
- 14 Timms, K. M. *et al.* Association of BRCA1/2 defects with genomic scores predictive of DNA damage repair deficiency among breast cancer subtypes. *Breast Cancer Res* **16**, 475, doi:10.1186/s13058-014-0475-x (2014).
- 15 Abkevich, V. *et al.* Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. *Br J Cancer* **107**, 1776-1782, doi:10.1038/bjc.2012.451 (2012).
- 16 Birkbak, N. J. *et al.* Telomeric allelic imbalance indicates defective DNA repair and sensitivity to DNA-damaging agents. *Cancer Discov* **2**, 366-375, doi:10.1158/2159-8290.Cd-11-0206 (2012).
- 17 Popova, T. *et al.* Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. *Cancer Res* **72**, 5454-5462, doi:10.1158/0008-5472.Can-12-1470 (2012).
- 18 Telli, M. L. *et al.* Homologous Recombination Deficiency (HRD) Score Predicts Response to Platinum-Containing Neoadjuvant Chemotherapy in Patients with Triple-Negative Breast Cancer. *Clin Cancer Res* **22**, 3764-3773, doi:10.1158/1078-0432.Ccr-15-2477 (2016).
- 19 Cancer Genome Atlas, N. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61-70, doi:10.1038/nature11412 (2012).
- 20 Shah, S. P. *et al.* The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* **486**, 395-399, doi:10.1038/nature10933 (2012).
- 21 Foulkes, W. D. *et al.* Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. *J Natl Cancer Inst* **95**, 1482-1485, doi:10.1093/jnci/djg050 (2003).
- 22 Gong, Y. *et al.* Metabolic-Pathway-Based Subtyping of Triple-Negative Breast Cancer Reveals Potential Therapeutic Targets. *Cell Metab* **33**, 51-64.e59, doi:10.1016/j.cmet.2020.10.012 (2021).
- 23 Jiang, Y. Z. *et al.* Genomic and Transcriptomic Landscape of Triple-Negative Breast Cancers: Subtypes and Treatment Strategies. *Cancer Cell* **35**, 428-440.e425, doi:10.1016/j.ccell.2019.02.001 (2019).
- 24 Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47, doi:10.1093/nar/gkv007 (2015).
- 25 Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287, doi:10.1089/omi.2011.0118 (2012).

- 26 Friedman, J., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw* **33**, 1-22 (2010).
- 27 Xiao, Y. *et al.* Multi-Omics Profiling Reveals Distinct Microenvironment Characterization and Suggests Immune Escape Mechanisms of Triple-Negative Breast Cancer. *Clin Cancer Res* **25**, 5002-5014, doi:10.1158/1078-0432.CCR-18-3524 (2019).
- 28 Tutt, A. & Ashworth, A. The relationship between the roles of BRCA genes in DNA repair and cancer predisposition. *Trends Mol Med* **8**, 571-576, doi:10.1016/s1471-4914(02)02434-6 (2002).
- 29 Sharma, P. *et al.* Impact of homologous recombination deficiency biomarkers on outcomes in patients with triple-negative breast cancer treated with adjuvant doxorubicin and cyclophosphamide (SWOG S9313). *Ann Oncol* **29**, 654-660, doi:10.1093/annonc/mdx821 (2018).
- 30 Mayerhoefer, M. E. *et al.* Introduction to Radiomics. *J Nucl Med* **61**, 488-495, doi:10.2967/jnumed.118.222893 (2020).
- 31 Liu, Z. *et al.* The Applications of Radiomics in Precision Diagnosis and Treatment of Oncology: Opportunities and Challenges. *Theranostics* **9**, 1303-1322, doi:10.7150/thno.30309 (2019).
- 32 Gillies, R. J., Kinahan, P. E. & Hricak, H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* **278**, 563-577, doi:10.1148/radiol.2015151169 (2016).
- 33 Aerts, H. J. *et al.* Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* **5**, 4006, doi:10.1038/ncomms5006 (2014).
- 34 Liu, Z. *et al.* Predicting distant metastasis and chemotherapy benefit in locally advanced rectal cancer. *Nat Commun* **11**, 4308, doi:10.1038/s41467-020-18162-9 (2020).
- 35 Huang, Y. Q. *et al.* Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer. *J Clin Oncol* **34**, 2157-2164, doi:10.1200/JCO.2015.65.9128 (2016).
- 36 Sun, R. *et al.* A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol* **19**, 1180-1191, doi:10.1016/S1470-2045(18)30413-3 (2018).
- 37 Trebeschi, S. *et al.* Predicting response to cancer immunotherapy using noninvasive radiomic biomarkers. *Ann Oncol* **30**, 998-1004, doi:10.1093/annonc/mdz108 (2019).
- 38 Farmer, H. *et al.* Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917-921, doi:10.1038/nature03445 (2005).

- 39 Robson, M. *et al.* Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. *N Engl J Med* **377**, 523-533, doi:10.1056/NEJMoa1706450 (2017).
- 40 Tutt, A. N. J. *et al.* Adjuvant Olaparib for Patients with BRCA1- or BRCA2-Mutated Breast Cancer. *N Engl J Med* **384**, 2394-2405, doi:10.1056/NEJMoa2105215 (2021).
- 41 Minckwitz, G. V. *et al.* Prediction of pathological complete response (pCR) by Homologous Recombination Deficiency (HRD) after carboplatin-containing neoadjuvant chemotherapy in patients with TNBC: Results from GeparSixto. *Journal of Clinical Oncology* **33**, 1004-1004, doi:10.1200/jco.2015.33.15\_suppl.1004 (2015).
- 42 Tutt, A. *et al.* Carboplatin in BRCA1/2-mutated and triple-negative breast cancer BRCAness subgroups: the TNT Trial. *Nat Med* **24**, 628-637, doi:10.1038/s41591-018-0009-7 (2018).
- 43 Jiang, Y. Z. *et al.* Molecular subtyping and genomic profiling expand precision medicine in refractory metastatic triple-negative breast cancer: the FUTURE trial. *Cell Res* **31**, 178-186, doi:10.1038/s41422-020-0375-9 (2021).
- 44 Gong, Y. *et al.* Metabolic-Pathway-Based Subtyping of Triple-Negative Breast Cancer Reveals Potential Therapeutic Targets. *Cell Metab* **33**, 51-64 e59, doi:10.1016/j.cmet.2020.10.012 (2021).

## Table 1

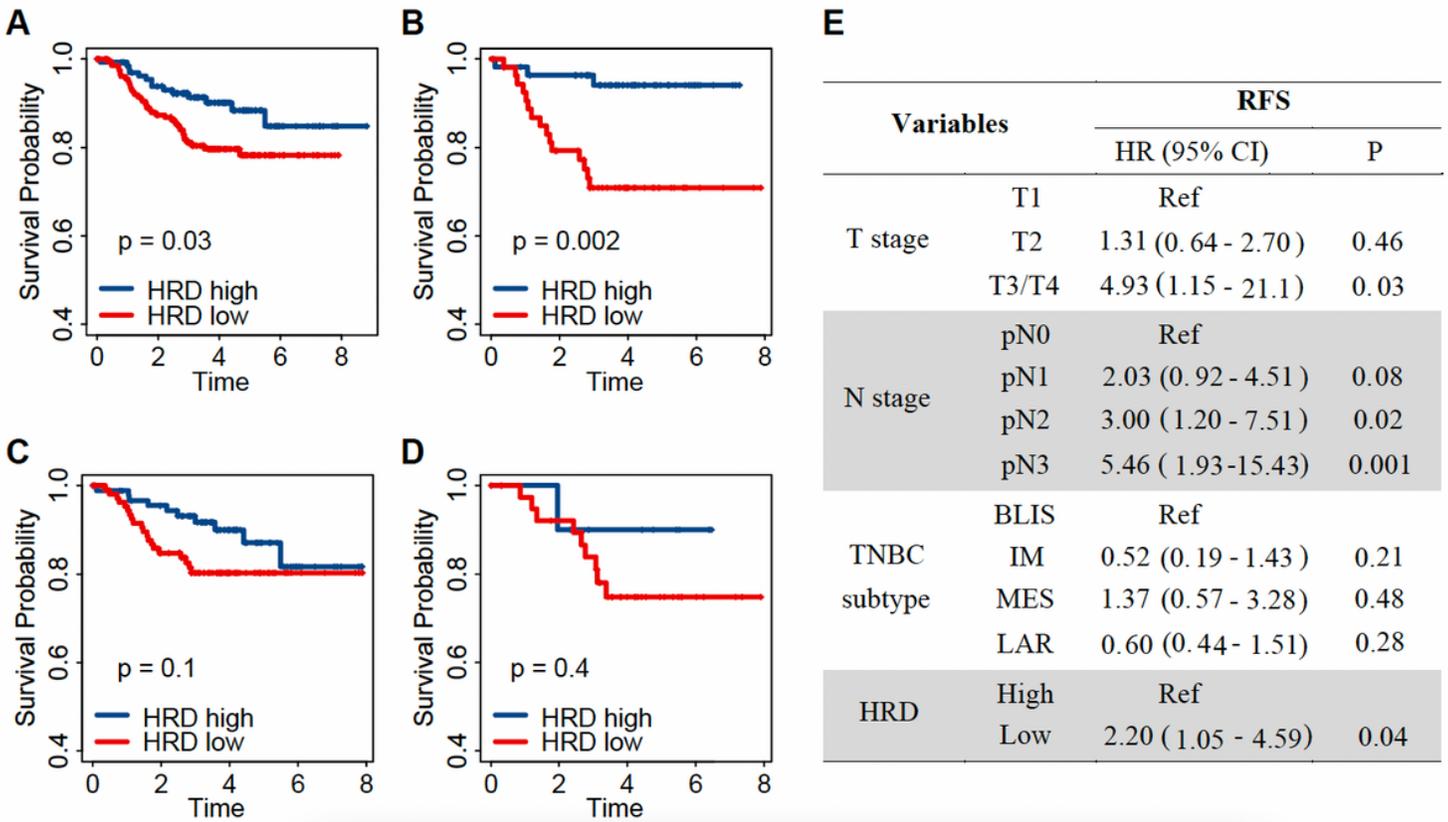
**Table 1 Clinical characteristics of high and low HRD TNBC patients**

	HRD high ( <i>n</i> = 135)		HRD low ( <i>n</i> = 208)		
Characteristics	No.	%	No.	%	P value
<b>Age (years)</b>					
Median	52		54		
IQR	44-58		46-62		
≤50	60	44.4%	85	40.9%	0.02
>50	75	55.6%	123	59.1%	
<b>Tumor size (cm)</b>					
≤2	43	31.9%	83	39.9%	0.29
>2,≤5	89	66.0%	119	57.2%	
>5	3	2.2%	6	2.9%	
<b>Positive LNs</b>					
0	92	68.2%	127	63.4%	0.10
1-3	31	23.0%	43	16.8%	
4-9	9	6.7%	20	8.9%	
≥10	3	2.2%	16	8.9%	
<b>Ki-67 (%)</b>					
≤20	12	8.9%	42	20.2%	0.0004
>20	119	88.1%	159	76.4%	
Unknown	4	3.2%	7	3.4%	

Abbreviations: HRD, homologous recombination deficiency; IQR, interquartile range; LN, lymph node.

## Figures

**Figure 1**

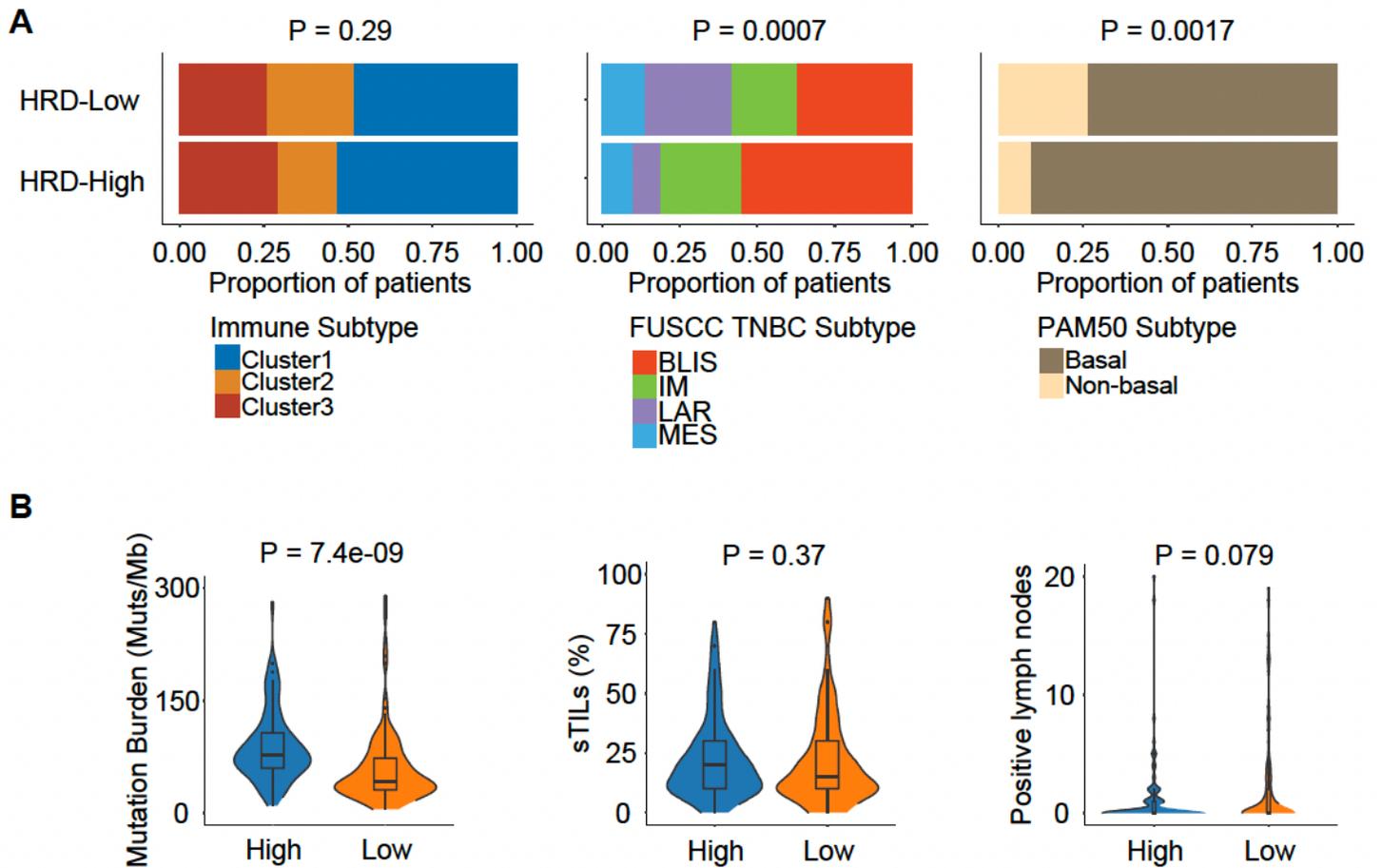


**Figure 1**

**Prognostic value of HRD in TNBC**

- (A) Probability of recurrence-free survival (RFS) in TNBC patients
- (B) Probability of RFS in BLIS subtype TNBC patients
- (C) Probability of RFS in Basal-like TNBC patients
- (D) Probability of RFS in non-Basal-like TNBC patients
- (E) Multivariate Cox proportional hazards model comprising T, N stages, TNBC subtypes and HRD score

**Figure 2**



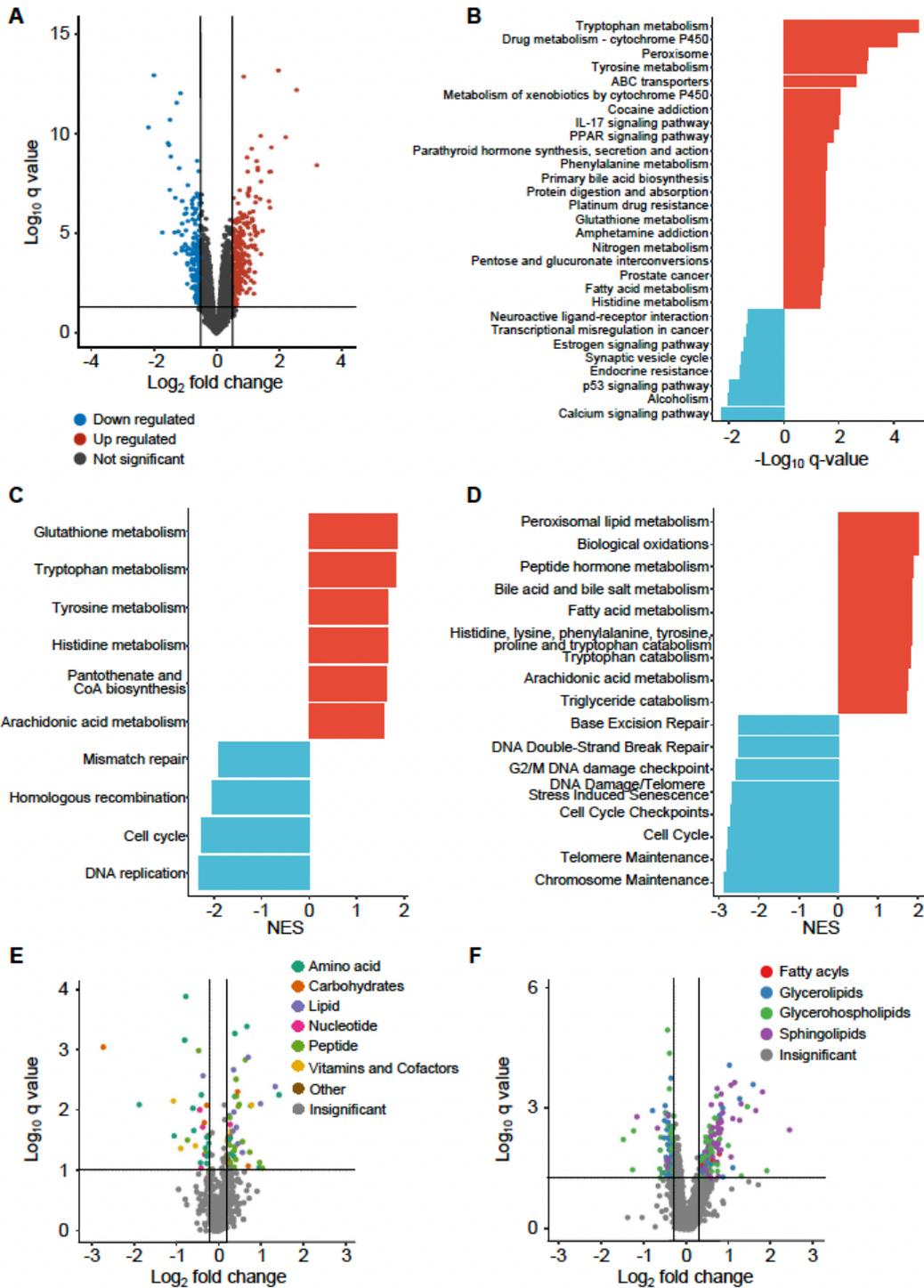
**Figure 2**

**Correlation analysis of HRD score and clinicopathological indices**

(A) The correlation between HRD score and immune subtype, FUSCC TNBC subtype and PAM50 subtype

(B) The correlation between HRD score and tumor mutation burden, stromal tumor infiltrating lymphocytes and positive lymph nodes

**Figure 3**



**Figure 3**

### Prognostic value of HRD in TNBC

(A) Differentially expressed genes (DEGs) between high and low HRD groups

(B) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis based on the DEGs

(C) Gene set enrichment analysis (GSEA) based on the KEGG database

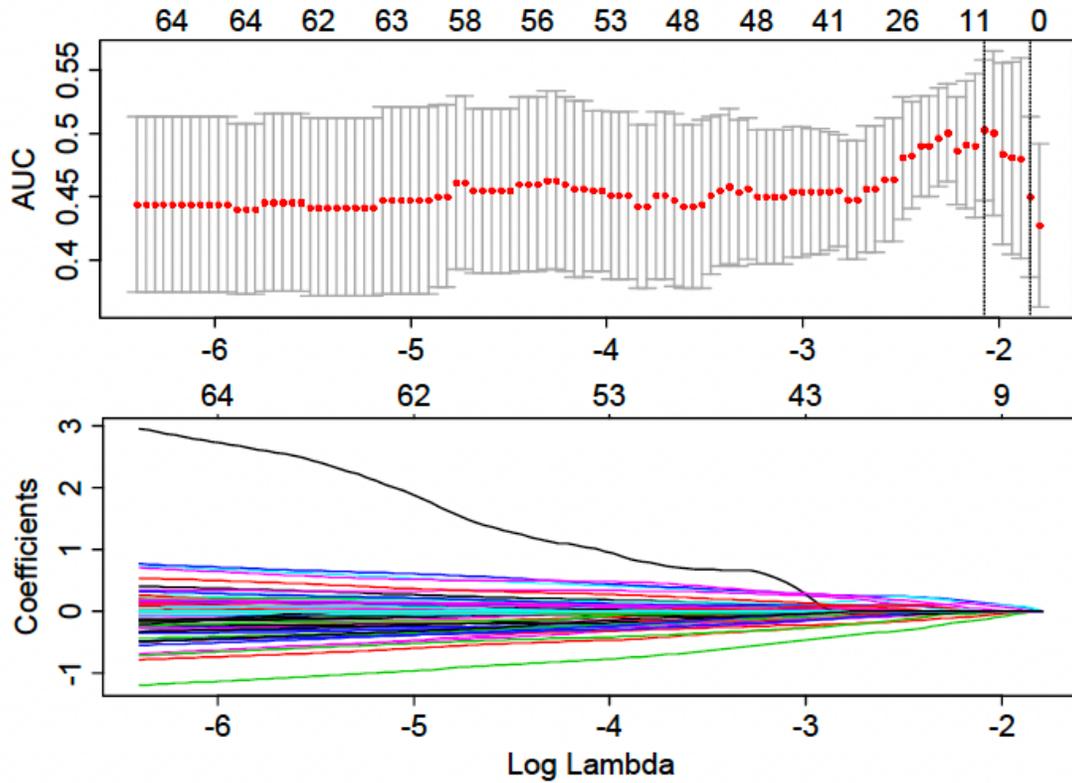
(D) GSEA based on the Reactome database

(E) Differentially expressed metabolites between high and low HRD groups

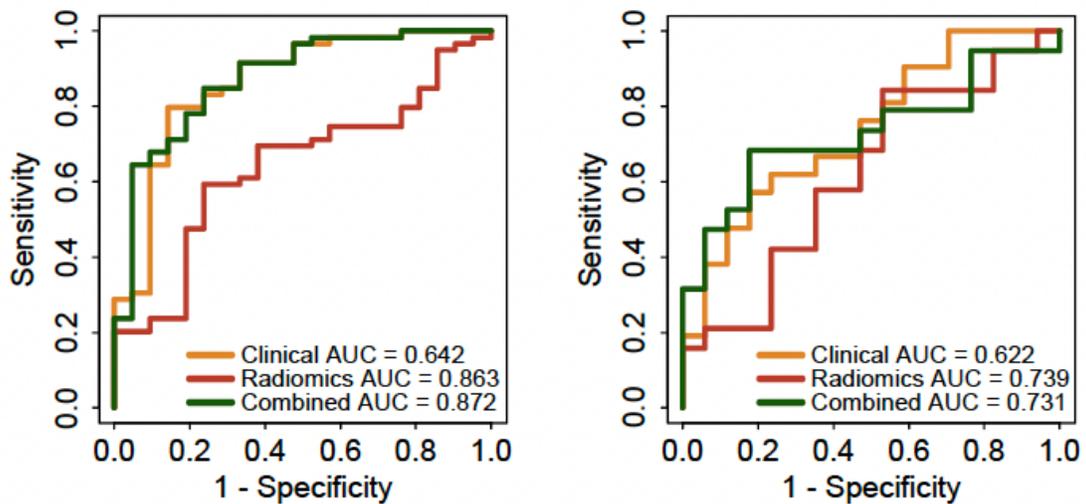
(F) Differentially expressed lipids between high and low HRD groups

**Figure 4**

**A**



**B**



**Figure 4**

## Performance of radiomics in HRD prediction

(A) Radiomics features selection process via the least absolute shrinkage and selection operator

(B) Performances of clinical model, radiomics model and combined model in the training set (left) and validation set (right)

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterials.pdf](#)