Title

A liquid biopsy signature for predicting early recurrence in patients with gastric cancer.

Journal

British Journal of Cancer, 128(6)

Authors

Okuno, Keisuke

Watanabe, Shuichi

Kanda, Mitsuro

et al.

Publication Date

2023-04-01

DOI

10.1038/s41416-022-02138-1

Peer reviewed

British Journal of Cancer www.nature.com/bjc

ARTICLE

Genetics and Genomics

A liquid biopsy signature for predicting early recurrence in patients with gastric cancer

Keisuke Okuno[1,2], Shuichi Watanabe[3], Souvick Roy[1], Mitsuro Kanda[4], Masanori Tokunaga[2], Yasuhiro Kodera[4], Yusuke Kinugasa[2] and Ajay Goel[1,5]✉

BACKGROUND: Gastric cancer (GC) patients who experience recurrence within the first year following surgery (early recurrence [ER]) exhibit worse prognosis. Herein, we established a microRNA-based liquid biopsy assay to predict ER in GC patients.
METHODS: A comprehensive biomarker discovery was performed by analysing miRNA expression profiling in 271 primary GC tumours. Thereafter, the expression of these biomarkers was validated in 290 GC cases, which included 218 tissues and 72 pre-treatment sera, from two independent institutions.
RESULTS: A panel of 8 miRNAs was identified during the initial biomarker discovery, and this panel could robustly predict ER in a tissue-based clinical cohort (area under the curve [AUC]: 0.81). Furthermore, a model combining the miRNA panel, microsatellite instability (MSI) status and tumour size exhibited superior predictive performance (AUC: 0.86), and was defined as a Prediction of Early Recurrence in GC (PERGC) signature, which was successfully validated in another independent cohort (AUC: 0.82). Finally, the PERGC signature was translated into a liquid biopsy assay (AUC: 0.81), and a multivariate regression analysis revealed this signature to be an independent predictor for ER (odds ratio: 11.20).
CONCLUSION: We successfully established a miRNA-based liquid biopsy signature that robustly predicts the risk of ER in GC patients.

British Journal of Cancer (2023) 128:1105–1116; https://doi.org/10.1038/s41416-022-02138-1

BACKGROUND

Gastric cancer (GC) is one of the major malignancies and the third leading cause of cancer-related deaths worldwide, with over 1,000,000 new cases and an estimated 768,793 deaths in 2020 [1, 2]. Recent progresses in treatment technologies have somewhat improved the prognosis in GC; however, 20–30% of patients with locoregional GC frequently develop tumour recurrence [3, 4]. In fact, tumour relapse is the leading cause of cancer-related deaths in patients who undergo curative gastrectomy, and ~40–50% of recurrences occur within the first year following surgery as reported in several retrospective studies [3, 4] and randomised controlled trials [5–8]. In this context, it is noteworthy that the subset of patients with early recurrence (ER; recurrence within one year after curative surgery) had a significantly poor prognosis, and their overall survival (OS) rates were worse compared to even those of stage IV GC patients [9–11]. This poor prognosis in patients with GC who experience ER might in part be attributed to poor compliance with chemotherapy following gastrectomy—an observation confirmed in the REGATTA trial, which demonstrated the superiority of chemotherapy alone vs. gastrectomy plus chemotherapy in advanced GC patients [12]. This study also illustrated that adequate chemotherapy compliance is in fact more important than radical gastrectomy in GC patients with distant metastases. This essentially suggests that such good chemotherapy compliance should be even more relevant in patients with ER, because such disease relapse following curative gastrectomy is supposed to be caused by the occult micro metastases at the time of resection [13–15]. In other words, unlike the current treatment strategy, systemic chemotherapy should be prioritised over radical gastrectomy in the treatment of patients with occult micro metastases. Hence, the identification of biomarkers that can accurately detect occult micro metastases and predict ER are of important clinical significance for the treatment of patients with GC.

While imaging and other tumour markers are sometimes used to detect recurrence in GC [10, 11, 15], these approaches have largely failed to detect occult micro-metastases following surgery. To meet this unmet clinical need, a variety of molecular biomarkers, including genes, microRNAs (miRNAs), circular RNAs, and DNA methylation alterations, have recently been developed to predict recurrence or ER in patients with GC [16–31]. However, due to their heavy reliance on surgically resected tissue-based approaches and inadequate accuracy, pathological TNM stage and

[1]Department of Molecular Diagnostics and Experimental Therapeutics, Beckman Research Institute of City of Hope, Biomedical Research Center, Monrovia, CA, USA. [2]Department of Gastrointestinal Surgery, Tokyo Medical and Dental University, Tokyo, Japan. [3]Department of Hepatobiliary and Pancreatic Surgery, Tokyo Medical and Dental University, Tokyo, Japan. [4]Department of Gastroenterological Surgery, Nagoya University Graduate School of Medicine, Nagoya, Japan. [5]City of Hope Comprehensive Cancer Center, Duarte, CA, USA. ✉email: ajgoel@coh.org

Received: 12 August 2022 Revised: 16 December 2022 Accepted: 22 December 2022
Published online: 11 January 2023

tumour histology remain as routine modality for predicting recurrence in GC [32, 33]. Since a liquid biopsy would enable a noninvasive and facile assay for predicting ER in GC patients, such an approach can offer an attractive solution for this important clinical problem.

The diagnostic and prognostic utility of serum or plasma-based miRNA levels have recently been highlighted, because their expression is frequently dysregulated in human cancer through various mechanisms. Furthermore, miRNAs have emerged as important molecular analytes for biomarker development as the expression of tumour-derived miRNAs is very stable in systemic circulation [34–39]. In addition, recent studies have also highlighted that microsatellite instability-high (MSI-H) status in GCs, which occurs in ~8–15% of cases, exhibits superior prognosis compared to those with MSI-low (MSI-L) or microsatellite stable (MSS) neoplasms [40–42]. In view of this evidence, we hypothesised that potentially a combination of miRNAs and the MSI status together might offer a more accurate prognostic model in patients with GC.

Herein, for the first time, we performed a genome-wide, systematic, and comprehensive analysis to discover a combination signature that includes novel miRNAs along with the MSI status for the prediction of ER in patients with stage II and III GC. This signature was first verified in a large, publicly available data set, followed by comprehensive validation in tissue specimens from two independent clinical cohorts. Finally, the performance of these biomarkers was translated into a liquid biopsy assay in pre-treatment blood specimens. Herein, we firstly report that a genomic signature comprising of circulating miRNA markers and tumoural MSI-status was robust in predicting ER in patients with GC.

METHODS

Comprehensive miRNA biomarker discovery

Firstly, stage II and III GC patients were categorised into 3 subgroups as follows: patients experienced recurrence within 1 year following surgery—the early recurrence (ER) subgroup; those who experienced recurrence after 1 year following surgery—the late recurrence (LR) subgroup; and those who never experienced recurrence—the no recurrence (NR) subgroup. The miRNA expression profiling of primary tumour tissues from The Cancer Genome Atlas (TCGA) were analysed to identify the miRNA biomarkers for the prediction of ER in patients with GC. TCGA data (level 3 miRNA-sequencing data) were downloaded from the University of California Santa Cruz Xena Browser (https://xenabrowser.net/). The data from a total of 352 stage II/III GC patients was present within the TCGA data set, and cases who had insufficient information about recurrence status were excluded. In total, miRNA expression profiling data from 271 GC patients (34 ER, 28 LR, and 209 NR) were analysed for the miRNA biomarker discovery.

To perform the cross-validation method in the biomarker discovery cohort, the 271 patients were randomly divided into a discovery training set (75% of patients, $n = 203$) and a discovery validation set (25% of patients, $n = 68$) [43]. In the discovery training set, the variable selection was performed by extracting the components that discriminated well between the groups using the partial least squares discriminant (PLS) method [44], followed by further variable selection and estimation of regression coefficients simultaneously using the least absolute shrinkage and selection operator (LASSO) analysis [43, 45]. Subsequently, the selected biomarkers were determined and evaluated using the receiver operating characteristics (ROC) curve and the area under the curve (AUC). Thereafter, the selected miRNA biomarkers panel were validated in the discovery validation set.

Patient cohorts

This study included analyses of 290 clinical specimens from patients with stage II/III disease enrolled at 2 independent institutions, which included 218 frozen tissue and 72 serum specimens. For the clinical biomarker training, frozen surgical tissue specimens from 124 stage II/III GC patients who were enrolled at the Tokyo Medical and Dental University, Japan, between 2009 and 2015 (clinical training cohort) were examined. For the clinical biomarker validation, another independent cohort of frozen surgical tissues from 94 stage II/III GC patients were enrolled at the Nagoya University, Japan, between 2010 and 2015 (clinical validation cohort) were analysed. Finally, for the performance evaluation in the liquid biopsy samples, pre-surgery serum specimens from 72 stage II/III GC patients (performance evaluation cohort) were examined. The serum specimens in the performance evaluation cohort were matched with the patient tissues, which were analysed within the clinical training and validation cohorts. The clinicopathological characteristics of each clinical cohort are shown in Supplementary Table S1.

Clinical data were collected from the electronic medical records and the clinical databases at each institution, and the data included patient demographics, comorbidities, recurrence-free survival (RFS), and OS. All tumours were histologically diagnosed as stage II/III GCs and were classified according to the Union for International Cancer Control (UICC) TNM classification of Malignant Tumours version 7. The treatment strategy was decided based on the Japanese GC treatment guideline [46]. Gastrectomy with D2 lymph node dissection was performed for clinically advanced-stage GC, while gastrectomy with D1 + lymph node dissection was performed for clinically early-stage GC. Patients received adjuvant chemotherapy with S-1 for 1 year [5, 47]. Follow-up was until death or for 5 years after radical surgery. A written informed consent was obtained from all patients. The study was conducted in compliance with the Declaration of Helsinki and was approved by the Institutional Review Boards of all participating institutions.

RNA extraction and real-time quantitative reverse-transcription polymerase chain reaction assays

Total RNA was extracted from frozen cancer tissue and pre-surgery serum specimens using an AllPrep DNA/RNA/miRNA Universal kit or the Qiagen miRNeasy Kit (Qiagen, Hilden, Germany), according to manufacturer's instructions. The total RNA was reverse transcribed to complementary DNA (cDNA) using a miRCURY LNA RT Kit (Qiagen) before real-time quantitative reverse-transcription polymerase chain reaction (qRT-PCR) assays. The qRT-PCR assays were examined using the QuantStudio 6 Flex RT-PCR System (Applied Biosystems, Foster City, CA, USA) and a SensiFAST SYBR Lo-ROX Kit (Bioline, London, UK). The miR-16-5p was used as an internal control, and $2^{-\Delta Ct}$ method was used for quantification. Normalised values were further $\log_{10}$ transformed [48]. The miRNA primers were purchased from Thermo Fisher Scientific, Waltham, MA, USA (Catalogue No: 4427975).

DNA extraction and MSI analysis

Total DNA was extracted from frozen cancer tissue using an AllPrep DNA/RNA/miRNA Universal kit (Qiagen), according to the manufacturer's instructions. MSI analysis was conducted using five mono-nucleotide repeat microsatellite markers (BAT-25, BAT-26, NR-21, NR-24, and NR-27) in a pentaplex PCR system, as described previously [49, 50].

Statistical analysis

All statistical analyses were performed using EZR [51], which is a graphical user interface for R (R Foundation for Statistical Computing, Vienna, Austria, version 4.0.3) designed to add statistical functions and is frequently used in biostatistics. Fisher's exact test was used to analyse categorical variables, and a two-sided Student's t test was used to analyse differences between continuous values. The cut-off points for continuous variables were divided by the mean value in each clinical cohort. Survival curves were constructed using the Kaplan–Meier method and were compared with the log-rank test. Binary logistic regression model was used to train a classifier based on the expression of eight miRNAs. ROC curve and AUC were used to evaluate the performance of the panel or signature or model for the prediction of ER. The factors acquired from univariate analysis ($P < 0.10$) were included in multivariate analysis with the binary logistic regression model. All P values were two-sided, and $P < 0.05$ was considered statistically significant.

RESULTS

GC patients with early recurrence exhibit significantly poor survival outcomes

At the outset, to confirm that GC patients who experienced ER exhibited poor prognosis, the OS rates were compared among patients within ER, LR, and NR subgroups, in two independent publicly available data sets (TCGA and GSE62254). The clinical information of

GSE62254 data set were downloaded from the Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo/). The clinicopathological characteristics of each public cohort are shown in Supplementary Table S2. The 3-year OS rates were significantly worse in ER vs. LR and NR subgroups in the (24.1% ER vs. 30.5% LR, $P < 0.01$; 24.1% ER vs. 77.5% NR, $P < 0.01$; Fig. 1a) and GSE62254 (3.3% ER vs. 56.8% LR, $P < 0.01$; 3.3% ER vs. 94.1% NR, $P < 0.01$; Fig. 1b). In addition, when comparing stage II/III patients who experienced ER vs. stage IV patients in these public cohorts, while the 3-year OS rates were similar among these subgroups in TCGA (24.1% vs. 26.4%; $P = 0.46$; Fig. 1c), the stage II/III patients with ER fared significantly worse in the GSE62254 data set (3.3% vs. 26.7%; $P = 0.04$; Fig. 1d). From these results, it was apparent that the stage II/III patients with early disease relapse exhibited worst prognosis among all GC patients. Furthermore, key clinicopathological factors, including age, sex, tumour location, Lauren classification, T stage, and lymph node metastasis, failed to predict ER and yielded AUC values ranging between 0.50 and 0.62 in both public data sets (Fig. 1e, f). These initial findings highlighted the need for the identification of biomarkers that can help predict ER in patients with GC for improving their survival outcomes.

Discovery of an 8-miRNA panel that predicts early recurrence in patients with GCs, which when combined with the MSI status further improved the overall predictive accuracy

The overall study design for the genome-wide expression profiling and discovery of miRNA biomarkers is illustrated in Fig. 2a. First, 271 GC patients within the TCGA data set were randomly divided into discovery training and validation sets. A panel of 8 candidate miRNAs were identified to predict ER using PLS [44] and LASSO analysis [43, 45] in the discovery training set (Fig. 2b), which revealed a promising predictive potential with a corresponding AUC value of 0.83 (95% confidence interval [CI]: 0.74–0.92; Fig. 2d). Subsequently, this 8-miRNA panel was validated in the discovery validation set, which yielded a comparable AUC value of 0.81 (95% CI: 0.69–0.93; Fig. 2e).

In TCGA data set, 51 of 271 stage II/III GC patients revealed MSI-H (18.8%), and ER subgroup notably had fewer number of MSI-H patients compared to LR and NR subgroups (5.9% ER vs. 20.7% LR vs. 20.5% NR; Fig. 2c). From these data, we hypothesised that MSI status might help improve the predictive capability of the 8-miRNA panel for the prediction of ER in GC patients. In support of our hypothesis, the combination of MSI status improved the AUC value of the 8-miRNA panel, from 0.83 to 0.87 (95% CI: 0.74–0.92; Fig. 2d) in the discovery training set, and from 0.81 to 0.87 (95% CI: 0.77–0.96; Fig. 2e) in the discovery validation set. Taken together, by analysing genome-wide expression profiling data, we identified a novel 8-miRNA biomarker panel for predicting ER in patients with stage II/III GC, which when combined together with the tumoural MSI status further improved the predictive accuracy for early disease relapse in patients with GC.

A genomic signature comprising of the 8-miRNA panel and MSI status successfully predicts early recurrence in patients with GC in clinical training cohort

To validate the prognostic outcomes for ER observed in the 2 public data sets, the OS rates were compared among GC patients with ER, LR, and NR. In line with the results from public data sets, the 5-year OS rates were significantly worse in the ER subgroup vs. LR and NR subgroups (6.2% ER vs. 44.3% LR, $P < 0.01$; 6.2% ER vs. 83.9% NR, $P < 0.01$; Fig. 3a). This clinical training cohort included 19 MSI-H patients (15.3%; Fig. 3b), and the ER subgroup included smaller number of MSI-H patients ($n = 2$, 8.0%) compared to LR ($n = 2$, 9.1%) and NR group ($n = 15$, 19.5%; Fig. 3c), which were consistent with the results in TCGA data set.

Next, we undertook qRT-PCR assays in the RNA derived from tissue specimens within the clinical training cohort, and using logistic regression analysis, a miRNA-based predictive panel for ER was established in the clinical training cohort. This panel revealed reasonable predictive potential, with an AUC value of 0.81 (95% CI: 0.73–0.90; $P < 0.01$; Fig. 3d). In terms of AUC values, the performance of this panel was superior to that of tumour size, which had the highest AUC value among all key clinicopathological factors, MSI status, and their combination (Fig. 3d). Consistent with our results in the biomarker discovery cohort, the combination of these data with the MSI status further improved the AUC value of our panel to 0.83 (95% CI: 0.74–0.91; $P < 0.01$; Fig. 3d). Furthermore, the inclusion of tumour size further enhanced the predictive potential of this signature, with a corresponding AUC value of 0.86 (95% CI: 0.77–0.93; $P < 0.01$; Fig. 3d). Taken together, the 8-miRNA panel, the MSI status, and the tumour size cumulatively was defined as a Prediction of Early Recurrence in GC (PERGC) signature, which robustly predicted ER in stage II/III GC patients. This PERGC signature was generated using multivariate logistic regression analyses, and the results of this analysis were shown in Supplementary Table S3. The formula of PERGC signature was as follows; $\mathrm{Logit}(P) = (1.6437 \times \text{miR-1294}) + (0.9122 \times \text{miR-378c}) + (1.2936 \times \text{miR-412-5p}) + (-0.0976 \times \text{miR-4749-5p}) + (2.5444 \times \text{miR-4306}) + (-0.2023 \times \text{miR-2114-5p}) + (-1.7655 \times \text{miR-6513-5p}) + (-1.9576 \times \text{miR-4677-3p}) + (-1.2803 \times \text{MSI status}) + (0.1284 \times \text{tumour size}) - 1.5861$. This PERGC signature was able to distinguish ER patients from those with LR and NR (ER vs. LR, $P < 0.01$; ER vs. NR, $P < 0.01$, respectively; Fig. 3e), and yielded a significantly greater predictive potential to predict ER vs. key clinicopathological factors (Fig. 3f). Collectively, a robust signature for the prediction of ER was established from a qRT-PCR-based miRNA biomarker assays, MSI analysis and key clinicopathological features in GC patients within the clinical training cohort.

The PERGC signature predicts early recurrence in patients with GC in the clinical validation cohort

Next, the biomarker validation of predictive potential for the PERGC signature was performed in an independent tissue-based cohort of patients with GC. In this clinical validation cohort, as was the case in the public data sets and the clinical training cohort, the 5-year OS rates were notably worse in the ER subgroup compared to the LR and NR subgroups (0.0% ER vs. 35.7% LR, $P = 0.04$; 0.0% ER vs. 90.1% NR, $P < 0.01$; Fig. 4a). In addition, the clinical validation cohort included 7 MSI-H patients (7.4%; Fig. 4b), and none of these patients experienced ER (Fig. 4c).

The PERGC signature, which was established in the clinical training cohort, demonstrated remarkable potential to predict ER even in the clinical validation cohort with an AUC value of 0.82 (95% CI: 0.72–0.91; $P < 0.01$; Fig. 4d, e). For the comparison of AUC values to key clinicopathological factors, this PERGC signature had a remarkably greater predictive capability for ER in the GC patients (Fig. 4f). In addition, subgroup analyses were performed to examine whether this PERGC signature was applicable to various subgroups of patients based upon age, histological subtypes of GC, and tumour stages. In this subgroup analyses, the PERGC signature yielded satisfactory performance for both older and younger patients, diffuse and intestinal types, as well as for stage II and III patients with GC (Fig. 4g). Overall, the PERGC model robustly allowed prediction of ER in GC patients even in the large independent tissue cohort of patients with GC.

Successful translation of the tissue-based PERGC model into a blood-based liquid biopsy assay

We next asked whether we can translate the tissue-based PERGC model into a non-invasive, blood-based liquid biopsy assay, which will allow for predicting ER in patient's blood prior to surgery. In order to answer this question, we obtained pre-surgery serum specimens from 72 stage II/III GC patients, who were matched with the tissue specimens that were analysed in the clinical training and validation cohorts. It was intriguing to observe that even in this performance evaluation cohort, consistent with the findings from tissue-based clinical cohorts, the 5-year OS rates

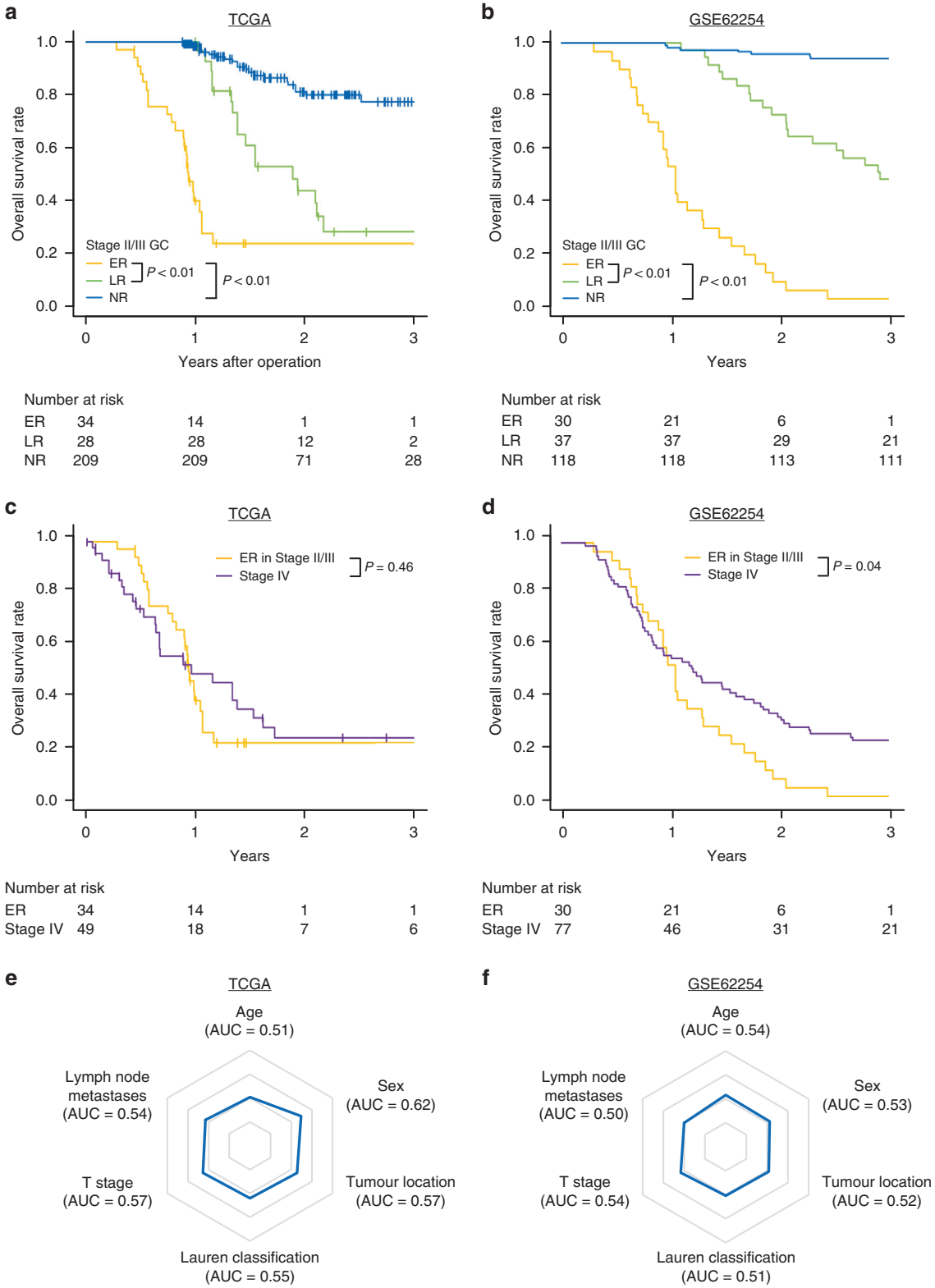


Fig. 1 Survival outcomes of stage II/III gastric cancer patients with early recurrence in two independent public data sets. a Kaplan-Meier curves for the overall survival for patients with ER ($n = 34$), LR ($n = 28$), and NR ($n = 209$) in TCGA data set ($n = 271$). **b** Kaplan–Meier curves for the overall survival for patients with ER ($n = 30$), LR ($n = 37$), and NR ($n = 118$) in GSE62254 data set ($n = 185$). **c** Kaplan–Meier curves for the overall survival in GC patients with ER in stage II/III ($n = 34$) and stage IV neoplasms ($n = 49$) in TCGA data set. **d** Kaplan–Meier curves for the overall survival in patients with ER in stage II/III ($n = 30$) and stage IV gastric cancer ($n = 77$) in GSE62254 data set. **e**, **f** Radar plot with AUC values for key clinicopathological factors to predict ER in TCGA (**e**) and GSE62254 (**f**) data sets. ER early recurrence, LR late recurrence, NR no recurrence, AUC area under the curve.

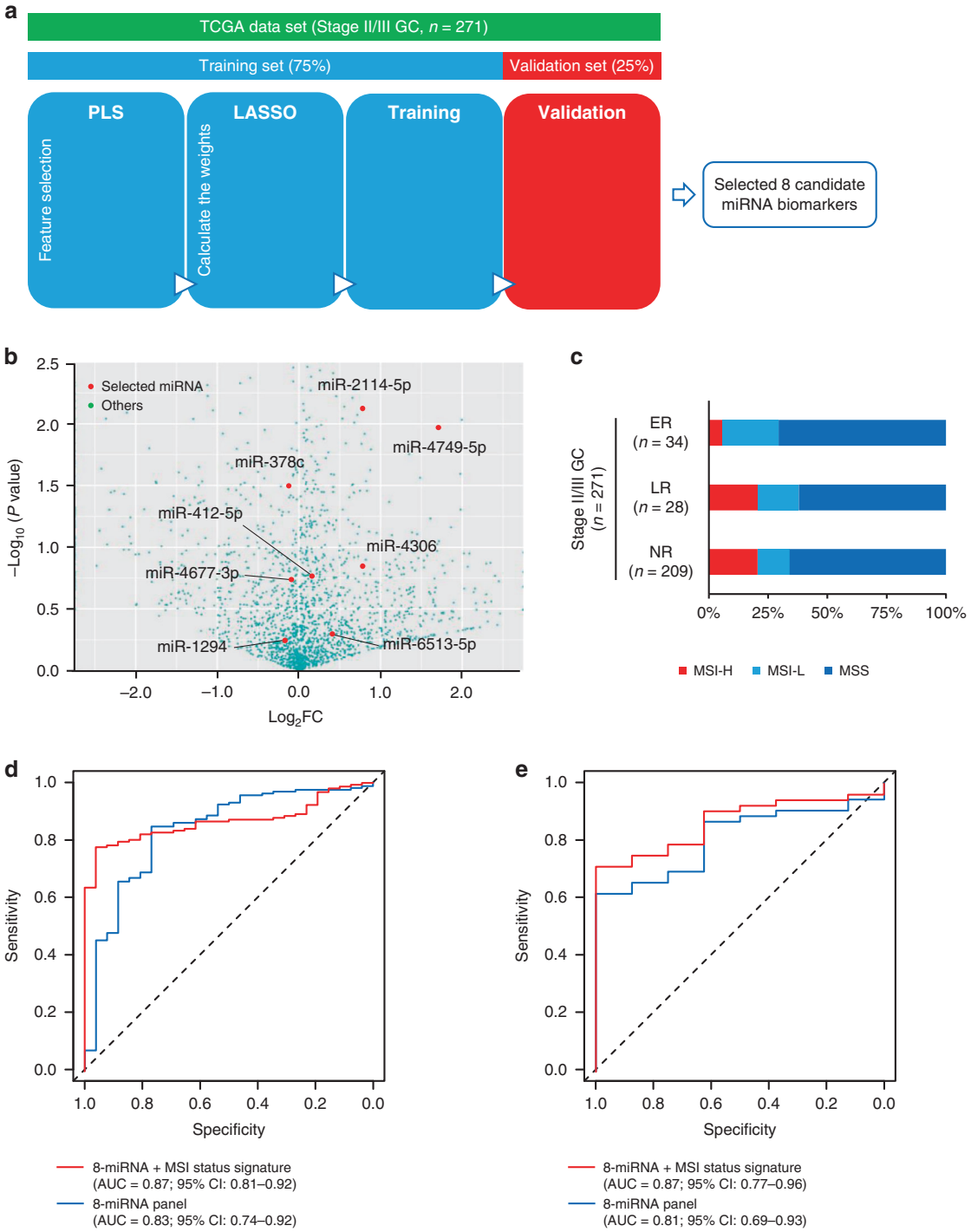
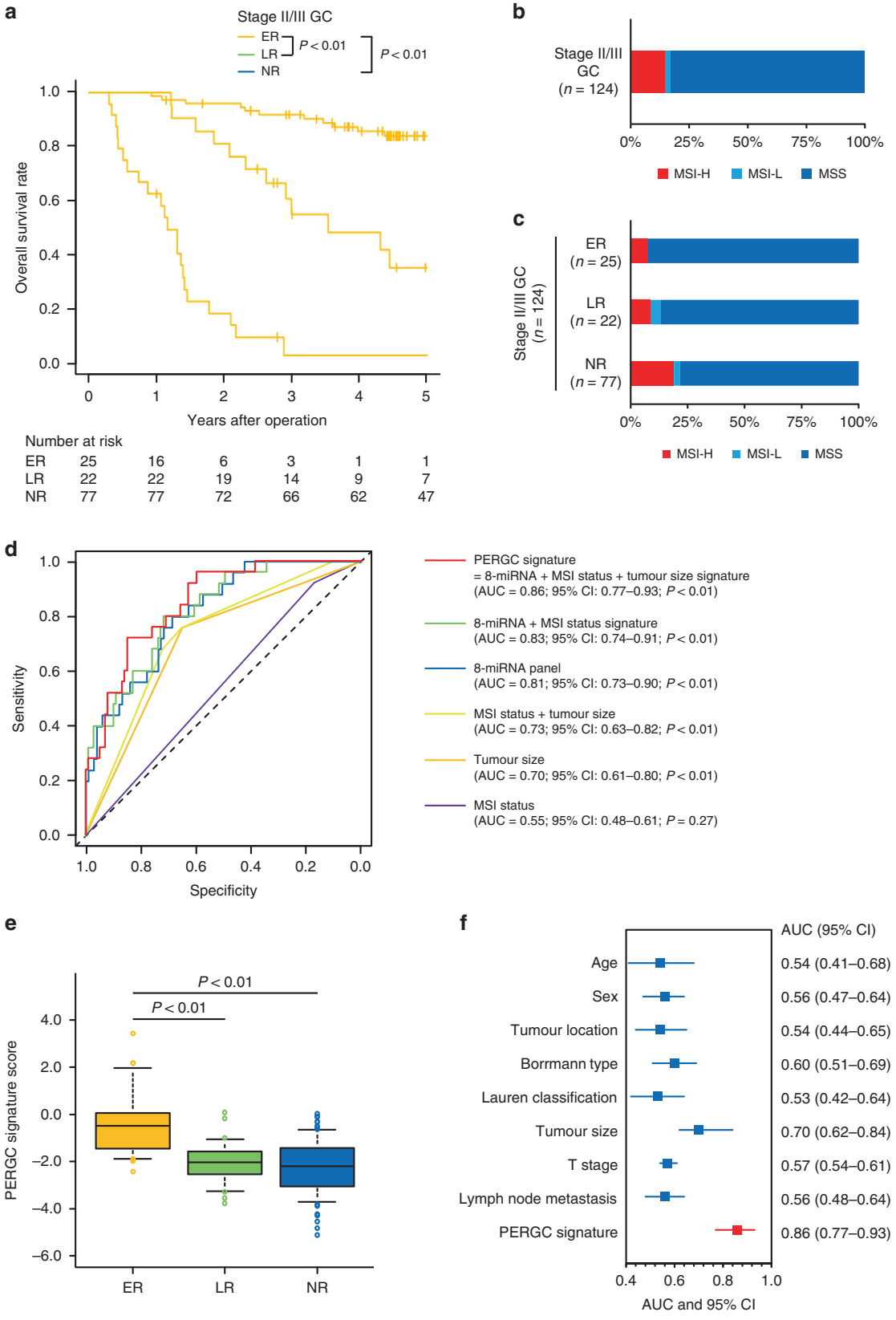


Fig. 2 Biomarker discovery phase of candidate miRNAs to predict early recurrence in primary tumours from gastric cancer patients. **a** Schematic of the study design for biomarker discovery in TCGA data set. **b** Volcano plot of 8 miRNAs to predict ER in primary tumours from gastric cancer patients in the discovery training set ($n = 203$). **c** Bar graph with MSI-H, MSI-L, and MSS patients in ER, LR and NR subgroups. **d** ROC curve values for 8-miRNA + MSI status signature or 8-miRNA panel alone in the discovery training set, wherein the miRNA + MSI status signature exhibited superior predictive accuracy in the discovery training set (AUC = 0.87 for the miRNA + MSI combination vs. AUC = 0.83 for the miRNA panel alone). **e** ROC curve values for 8-miRNA + MSI status signature or 8-miRNA panel in the discovery validation set, where the miRNA + MSI status signature exhibited superior performance in the discovery validation set (AUC = 0.87 for miRNA + MSI combination vs. AUC = 0.81 for the miRNA panel). GC gastric cancer, PLS partial least squares discriminant analysis, LASSO least absolute shrinkage and selection operator analysis, FC fold change, MSI-H microsatellite instability high, MSI-L microsatellite instability low, MSS microsatellite stability, ER early recurrence, LR late recurrence, NR no recurrence, ROC receiver operating characteristics, AUC area under the curve, CI confidence interval.



were significantly worse in ER subgroup compared to patients with LR or NR (0.0% ER vs. 33.3% LR, $P = 0.04$; 0.0% ER vs. 85.0% NR, $P < 0.01$; Fig. 5a). For the MSI status, performance evaluation cohort involved 7 MSI-H patients (9.7%; Fig. 5b), but the ER subgroup did not include any MSI-H patients, unlike LR and NR subgroups (Fig. 5c).

In the pre-surgery serum specimens, the miRNA-based panel showed reasonable predictive potential, with an AUC value of 0.76

Fig. 3 Clinical training phase of the PERGC signature for predicting the early recurrence in tissue specimens from stage II/III GC patients. **a** Kaplan–Meier curves for the overall survival in patients with ER ($n = 25$), LR ($n = 22$), and NR ($n = 77$) in the clinical training cohort ($n = 124$). **b** Bar graph with MSI-H, MSI-L, and MSS patients in the clinical training cohort ($n = 124$). **c** Bar graph with MSI-H, MSI-L, and MSS patients in ER, LR and NR subgroups within the clinical training cohort. **d** ROC curve values for PERGC signature, 8-miRNA + MSI status signature, 8-miRNA panel alone, MSI status + tumour size, tumour size, and MSI status to predict ER in the clinical training cohort. **e** Box plot for the PERGC signature scores with ER, LR and NR. **f** Forest plot with AUC values for clinicopathological factors and PERGC signature for the prediction of ER in the clinical training cohort. GC gastric cancer, MSI-H microsatellite instability high, MSI-L microsatellite instability low, MSS microsatellite stability, ER early recurrence, LR late recurrence, NR no recurrence, ROC receiver operating characteristics, AUC area under the curve, CI confidence interval.

(95% CI: 0.62–0.90; $P < 0.01$) and exhibited superior performance for predicting ER when combined together with the MSI status and tumour size as evidenced by a higher AUC value of 0.81 (95% CI: 0.70–0.93; $P < 0.01$; Fig. 5d). The PERGC model exhibited a superior predictive accuracy compared to key clinicopathological factors in terms of AUC value (Fig. 5e). For the sensitivity to predict ER, several clinicopathological factors had higher sensitivity compared to our model, however, their specificity was extremely low, with <30.0%, while the same tendency was observed for the clinicopathological factors with high specificity (Supplementary Table S4). Compared to clinicopathological factors, our model revealed acceptable overall performance in terms of sensitivity and specificity, which was 83.3% and 71.7%, respectively. In addition, according to multivariate logistic regression analysis, which included factors derived from the univariate analysis ($P < 0.10$), the PERGC model was identified as only one significant independent predictor for ER in GC patients (OR: 11.20; 95% CI: 2.07–60.80; $P < 0.01$; Table 1). Furthermore, when categorising all patients into high and low PERGC model score groups using cut-off thresholds derived from Youden's index [52], the 5-year RFS rates were worse in patients with PERGC-high score vs. those with PERGC-Low score (50.9% vs. 70.9%; $P = 0.03$; Fig. 5f). Interestingly, not only the 5-year RFS but also the 5-year OS rates were worse in patients with high vs. low scores (42.1% vs. 73.4%; $P = 0.01$; Fig. 5g). Collectively, the PERGC model was successfully translated into a liquid biopsy assay in pre-surgery serum specimens, which when combined with the tumoural MSI status and other clinical factors resulted in a significantly superior prediction for ER in patients with stage II/III GCs.

DISCUSSION

Current development of treatment technologies and strategies has largely improved the prognosis of patients with GC. Unfortunately, however, approximately 30% of patients with locoregional lesions develop tumour recurrence [3, 4], and the median survival following recurrence in these patients is relatively short—often shorter than 1 year [53, 54]. In particular, among patients with recurrence following surgery, those with early disease relapse exhibit worse prognosis, and are considered as subpopulation of patients with GC with the worst OS outcomes [9–11]. In view of this evidence, patients with ER must be treated and managed differently for improving their survival; however, no clinicopathological factors and molecular biomarkers currently available are adept for clinical application in identifying these patients with higher rates for ER in gastric neoplasia. Therefore, a robust pre-treatment prediction of patients who have a higher risk for developing ER using molecular biomarkers is an important clinical unmet need, which must be addressed for achieving precision treatment strategies for patients with GC. In the present study, we identified a panel of 8-miRNAs for predicting ER; which when combined with the tumoural MSI status and key clinical factors further improved the predictive accuracy of these markers. Following initial discovery, we were able to successfully validate the performance of these markers in two independent large tissue-based clinical cohorts. Most importantly, we were able to successfully translate this tissue-based molecular assay into a blood-based liquid biopsy assay, which ensures a noninvasive and facile assay for predicting ER, by analysing matched pre-surgery serum specimens of GC patients.

Numerous studies have previously identified various molecular biomarkers, including genes, various non-coding RNAs, DNA methylation alterations, and somatic copy-number alteration, as the significant predictors of tumour recurrence in GC patients. In particular, for the early recurrence, several clinicopathological factors and molecular biomarkers, which include miRNAs and circular RNAs, have been reported as the predictors of ER in GC patients [10, 11, 19, 21, 22]. However, due to a variety of limitations of these studies, none of these biomarkers have translated into the clinical practice. Our present study has several advantages compared to these previous studies. First, a genome-wide miRNA-based biomarker discovery was performed using a large, publicly available data set, which was able to screen perhaps the largest number of miRNAs and led to the identification of potentially more robust biomarkers. Second, our model was comprehensively validated in independent large clinical cohorts of patients from multiple institutions, which ensured generalisability of these markers. Third, the PERGC model was successfully translated into a blood-based liquid biopsy assay, which offers simplicity, noninvasiveness and cost-effectiveness. A liquid biopsy assay using pre-surgery specimens will inform the clinicians prior to the initiation of treatment about the probability of a patient to experience ER, which will allow precision and tailor-made treatment planning for such high-risk individuals. Fourth, our study had new interesting findings that ER subgroup included relatively smaller number of MSI-H patients compared to LR and NR subgroups in both TCGA and multiple clinical cohorts. Unfortunately, MSI status alone had insufficient potential to predict ER (Supplementary Fig. S1); however, our findings demonstrated the combination of MSI status and molecular biomarkers revealed great potential to predict ER in GC patients.

In the present study, we defined early recurrence as an event within 1 year after curative surgery for GC. According to the Japanese GC treatment guidelines [5, 46, 47], most of patients in this study received 1-year adjuvant chemotherapy with S-1 following curative resection of tumours, indicating that early recurrence occurred during the course of these adjuvant chemotherapies. In other words, ER group in our study was likely a subset of non-responders to adjuvant chemotherapy treatment. Collectively, our model might possibly identify patients who are ineffective to these therapies, which also can provide the fundamentals for precision medicine in an adjuvant chemotherapy setting for GC.

To clarify the biological relevance of our miRNA biomarkers, the downstream target gene and enrichment pathway analyses were performed using the miRDB [55, 56] and DAVID bioinformatic databases [57]. The target genes of 8 miRNA biomarkers (target score > 80) were significantly enriched (fold enrichment >2.0 and $P < 0.01$) in GC (Supplementary Fig. S2). Of note, enrichment pathway analyses identified multiple cancer-related pathways, such as Wnt, Hippo, mTOR, cAMP, and Rap1 signalling pathway (Supplementary Fig. S2), highlighting the biological relevance of our miRNA biomarkers in the prediction of ER in patients with GC.

We would like to acknowledge potential limitations of this study. First, this was a retrospective cohort study and analysed a

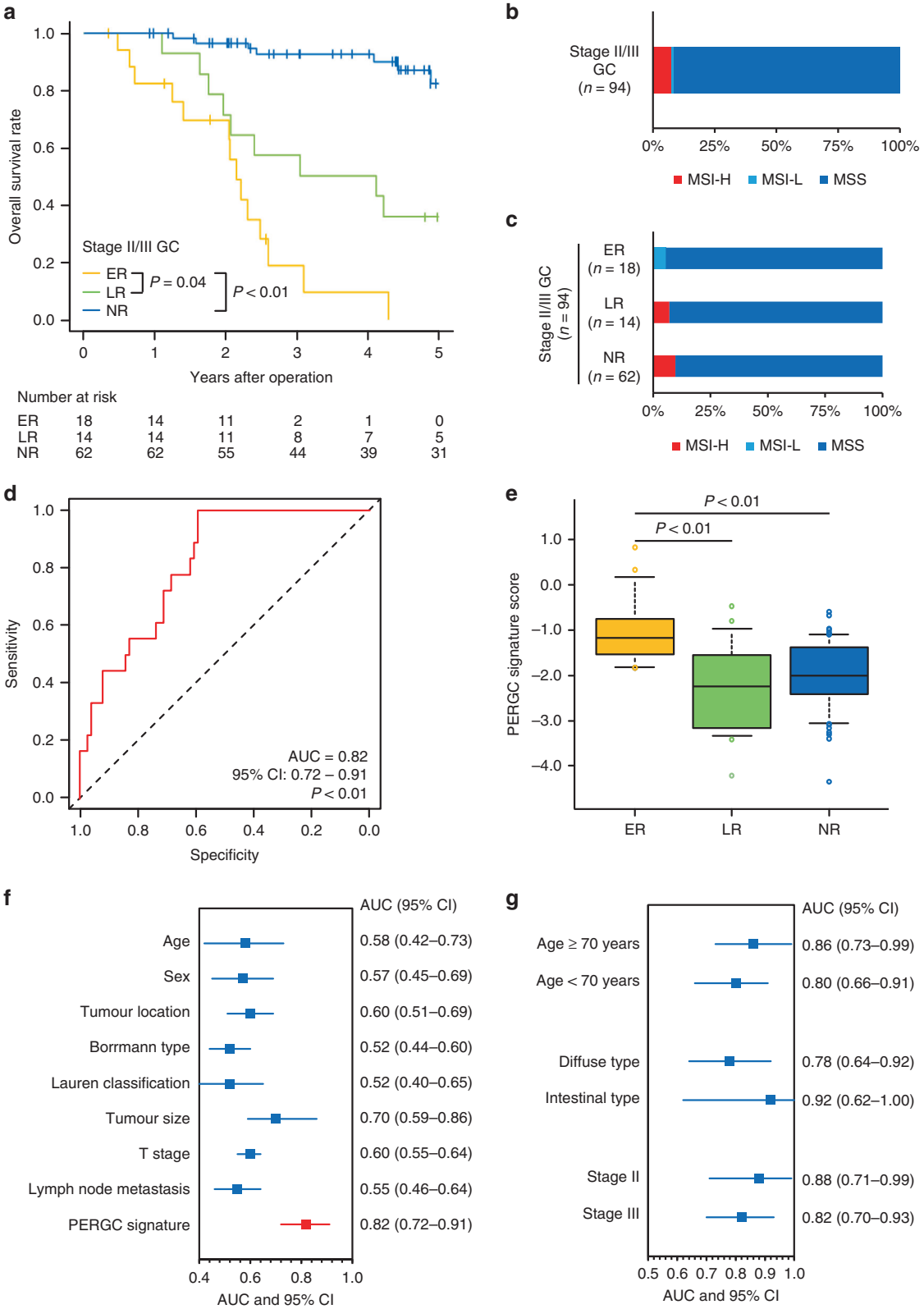
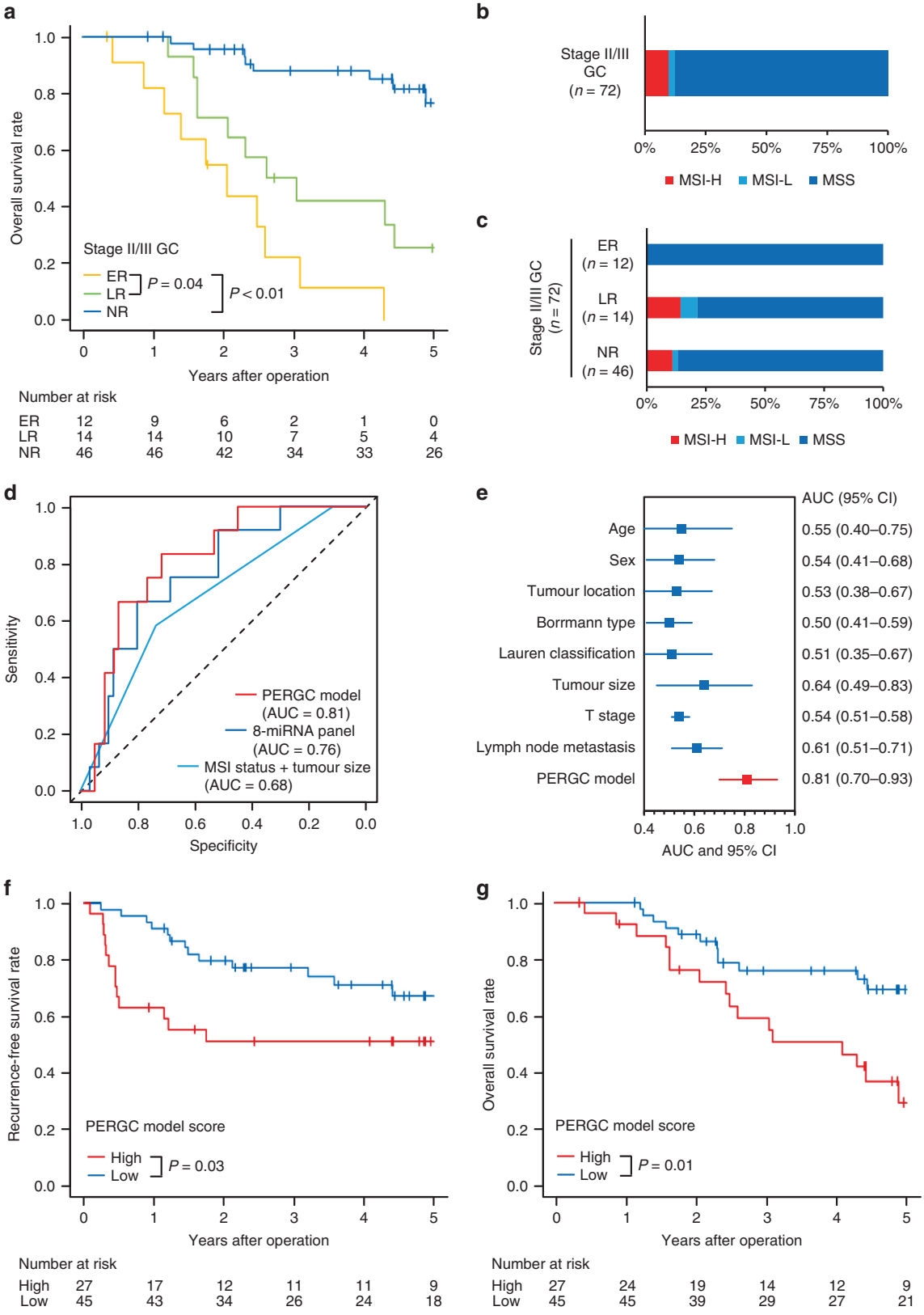


Fig. 4 Clinical validation phase for the PERGC signature in predicting early recurrence in tissue specimens from stage II/III GC patients. **a** Kaplan–Meier curves for the overall survival in patients with ER ($n = 18$), LR ($n = 14$), and NR ($n = 62$) in the clinical validation cohort ($n = 94$). **b** Bar graph with MSI-H, MSI-L, and MSS patients in the clinical validation cohort ($n = 94$). **c** Bar graph with MSI-H, MSI-L, and MSS patients in ER, LR and NR subgroups within the clinical validation cohort. **d** ROC curve values for the PERGC signature to predict ER in the clinical validation cohort (AUC = 0.82). **e** Box plot for the PERGC signature scores for the GC patients with ER, LR, and NR. **f** Forest plot with AUC values of key clinicopathological factors and PERGC signature for the prediction of ER in the clinical validation cohort. **g** Forest plot with AUC values of the PERGC signature in subgroup analyses. GC gastric cancer, MSI-H microsatellite instability high, MSI-L microsatellite instability low, MSS microsatellite stability, ER early recurrence, LR late recurrence, NR no recurrence, ROC receiver operating characteristics, AUC area under the curve, CI confidence interval.



rather modest number of patients from Asian cohorts. In addition, although prognostic biomarkers for GC should be considered for evaluation separately in various subgroups of patients based on the status of MSI, Epstein-Barr virus infection, and human epidermal growth factor receptor 2, due to the limited number of patients in our article such analyses were not possible and should be appropriately addressed in future studies. Second, the PERGC model was successfully translated into a liquid biopsy assay in the present study, however, our model was not validated in another independent blood-based cohort. Therefore, for the

Fig. 5 Clinical performance evaluation phase for the PERGC model as a liquid biopsy assay for predicting early recurrence in pre-surgery serum specimens from stage II/III GC patients. a Kaplan–Meier curves for the overall survival in patients with ER ($n = 12$), LR ($n = 14$), and NR ($n = 46$) in the performance evaluation cohort ($n = 72$). **b** Bar graph with MSI-H, MSI-L, and MSS patients in the performance evaluation cohort ($n = 72$). **c** Bar graph with MSI-H, MSI-L, and MSS patients in ER, LR and NR subgroups in the performance evaluation cohort. **d** ROC curve values for the PERGC model, 8-miRNAs panel, and MSI status + tumour size to predict ER in performance evaluation cohort. **e** Forest plot with AUC values for each of the clinicopathological factors and PERGC model for the prediction of ER in the performance evaluation cohort. **f** Kaplan–Meier curves for the recurrence-free survival in patients with PERGC-high ($n = 27$) and PERGC-low scores ($n = 45$). **g** Kaplan–Meier curves illustrating the overall survival in patients with PERGC-high ($n = 27$) and PERGC-low scores ($n = 45$). GC gastric cancer, MSI-H microsatellite instability high, MSI-L microsatellite instability low, MSS microsatellite stability, ER early recurrence, LR late recurrence, NR no recurrence, ROC receiver operating characteristics, AUC area under the curve, CI confidence interval.

Table 1. Univariate and multivariate analysis for key clinicopathological and molecular markers contributing towards early recurrence in stage II/III patients with gastric cancer.

	Univariate		Multivariate	
	OR (95% CI)	***P* value**	**OR (95% CI)**	***P* value**
Age, >68 versus ≤68 years	1.50 (0.43–5.24)	0.53		
Sex, female versus male	1.67 (0.38–7.27)	0.50		
Tumour location, upper versus middle and lower	1.29 (0.31–5.32)	0.73		
Borrmann type, type 4 versus others	1.00 (0.11–9.42)	1.00		
Lauren classification, diffuse versus intestinal	1.07 (0.31–3.69)	0.92		
Tumour size, >6.6 versus ≤6.6 cm	3.02 (0.85–10.80)	0.09	1.40 (0.33–5.89)	0.64
Clinical T stage, T3–4 versus T1–2	1.20 (0.32–5.20)	0.99		
Lymph node metastases, positive versus negative	4.71 (0.57–39.30)	0.15		
PERGC model score, high versus low	12.60 (2.51–63.80)	**<0.01**	11.20 (2.07–60.80)	**<0.01**

Bold values indicate statistical significance $p < 0.05$.
OR odds ratio, *CI* confidence interval.

further confirmation of this model, these results should be further validated in future large prospective clinical trials. Third, the clinical cohorts in this study did not include GC patients who received neoadjuvant treatment (NAT). The NAT for GC was not the standard treatment in Japan at the time of patient enrolment for our study [46]; while locally advanced GC patients in the western countries often receive neoadjuvant chemotherapy, such as FLOT4 [58] or FP [59]. Therefore, further validation of the biomarkers reported in our study in patients who received NAT in future studies might further bolster their clinical significance in clinical practice. Fourth, although the PERGC model was successfully translated into a liquid biopsy assay, such an assay might not represent complete transcriptomic profile of cancer tissues; hence, future genome-wide profiling in liquid biopsy specimens might reveal additional miRNA biomarkers for predicting ER. Despite these limitations, our stringent miRNA-based liquid biopsy assay offers a potential opportunity to develop a precision medicine approach for the identification of high-risk GC patients that are likely to manifest early disease relapse; and will allow an improved management of these patients by offering them appropriate treatments to mitigate their chances for experiencing ER.

In conclusion, we have successfully identified and established a miRNA-based liquid biopsy assay combined with tumoural MSI status, that allows robust prediction of ER in GC patients. These findings can provide the fundamentals for developing precision medicine approaches for identifying high-risk GC patients that are likely to develop ER in pre-operative settings, so that they can be offered alternate treatment options for reducing their risk for recurrence and improving their survival.

DATA AVAILABILITY

The data sets analysed during the current study are available from the corresponding author on reasonable request.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71:209–49.
2. Arnold M, Abnet CC, Neale RE, Vignat J, Giovannucci EL, McGlynn KA, et al. Global burden of 5 major types of gastrointestinal cancer. Gastroenterology 2020;159:335.e15–49.e15.
3. Alnoor M, Boys JA, Worrell SG, Oh DS, Hagen JA, DeMeester SR. Timing and pattern of recurrence after gastrectomy for adenocarcinoma. Am Surg. 2015;81:1057–60.
4. Spolverato G, Ejaz A, Kim Y, Squires MH, Poultsides GA, Fields RC, et al. Rates and patterns of recurrence after curative intent resection for gastric cancer: a United States multi-institutional analysis. J Am Coll Surg. 2014;219:664–75.
5. Sasako M, Sakuramoto S, Katai H, Kinoshita T, Furukawa H, Yamaguchi T, et al. Five-year outcomes of a randomized phase III trial comparing adjuvant chemotherapy with S-1 versus surgery alone in stage II or III gastric cancer. J Clin Oncol. 2011;29:4387–93.
6. Yoshida K, Kodera Y, Kochi M, Ichikawa W, Kakeji Y, Sano T, et al. Addition of docetaxel to oral fluoropyrimidine improves efficacy in patients with stage III gastric cancer: interim analysis of JACCRO GC-07, a randomized controlled trial. J Clin Oncol. 2019;37:1296–304.
7. Noh SH, Park SR, Yang HK, Chung HC, Chung IJ, Kim SW, et al. Adjuvant capecitabine plus oxaliplatin for gastric cancer after D2 gastrectomy (CLASSIC): 5-year follow-up of an open-label, randomised phase 3 trial. Lancet Oncol. 2014;15:1389–96.
8. Di Costanzo F, Gasperoni S, Manzione L, Bisagni G, Labianca R, Bravi S, et al. Adjuvant chemotherapy in completely resected gastric cancer: a randomized phase III trial conducted by GOIRC. J Natl Cancer Inst. 2008;100:388–98.
9. Katai H, Ishikawa T, Akazawa K, Isobe Y, Miyashiro I, Oda I, et al. Five-year survival analysis of surgically resected gastric cancer cases in Japan: a retrospective analysis of more than 100,000 patients from the nationwide registry of the Japanese Gastric Cancer Association (2001-2007). Gastric Cancer. 2018;21:144–54.
10. Chiang CY, Huang KH, Fang WL, Wu CW, Chen JH, Lo SS, et al. Factors associated with recurrence within 2 years after curative surgery for gastric adenocarcinoma. World J Surg. 2011;35:2472–8.
11. Kang WM, Meng QB, Yu JC, Ma ZQ, Li ZT. Factors associated with early recurrence after curative surgery for gastric cancer. World J Gastroenterol. 2015;21:5934–40.
12. Fujitani K, Yang HK, Mizusawa J, Kim YW, Terashima M, Han SU, et al. Gastrectomy plus chemotherapy versus chemotherapy alone for advanced gastric cancer with a single non-curable factor (REGATTA): a phase 3, randomised controlled trial. Lancet Oncol. 2016;17:309–18.

13. Van Cutsem E, Sagaert X, Topal B, Haustermans K, Prenen H. Gastric cancer. Lancet. 2016;388:2654–64.
14. Bickenbach KA, Gonen M, Strong V, Brennan MF, Coit DG. Association of positive transection margins with gastric cancer survival and local recurrence. Ann Surg Oncol. 2013;20:2663–8.
15. Marrelli D, De Stefano A, de Manzoni G, Morgagni P, Di Leo A, Roviello F. Prediction of recurrence after radical surgery for gastric cancer: a scoring system obtained from a prospective multicenter study. Ann Surg. 2005;241:247–55.
16. Lee KW, Lee SS, Hwang JE, Jang HJ, Lee HS, Oh SC, et al. Development and validation of a six-gene recurrence risk score assay for gastric cancer. Clin Cancer Res. 2016;22:6228–35.
17. Motoori M, Takemasa I, Yano M, Saito S, Miyata H, Takiguchi S, et al. Prediction of recurrence in advanced gastric cancer patients after curative resection by gene expression profiling. Int J Cancer. 2005;114:963–8.
18. Yan Z, Xiong Y, Xu W, Li M, Cheng Y, Chen F, et al. Identification of recurrence-related genes by integrating microRNA and gene expression profiling of gastric cancer. Int J Oncol. 2012;41:2166–74.
19. Liu X, Zhang X, Zhang Z, Chang J, Wang Z, Wu Z, et al. Plasma microRNA-based signatures to predict 3-year postoperative recurrence risk for stage II and III gastric cancer. Int J Cancer. 2017;141:2093–102.
20. Yang Y, Qu A, Zhao R, Hua M, Zhang X, Dong Z, et al. Genome-wide identification of a novel miRNA-based signature to predict recurrence in patients with gastric cancer. Mol Oncol. 2018;12:2072–84.
21. Zhang Y, Li J, Yu J, Liu H, Shen Z, Ye G, et al. Circular RNAs signature predicts the early recurrence of stage III gastric cancer after radical surgery. Oncotarget. 2017;8:22936–43.
22. Imaoka H, Toiyama Y, Okigami M, Yasuda H, Saigusa S, Ohi M, et al. Circulating microRNA-203 predicts metastases, early recurrence, and poor prognosis in human gastric cancer. Gastric Cancer. 2016;19:744–53.
23. Goel A. Somatostatin methylation as a biomarker for gastric cancer: ready for "prime time" or for further validation? Dig Dis Sci. 2011;56:1–3.
24. Okugawa Y, Toiyama Y, Hur K, Toden S, Saigusa S, Tanaka K, et al. Metastasis-associated long non-coding RNA drives gastric cancer development and promotes peritoneal metastasis. Carcinogenesis. 2014;35:2731–9.
25. Toiyama Y, Okugawa Y, Goel A. DNA methylation and microRNA biomarkers for noninvasive detection of gastric and colorectal cancer. Biochem Biophys Res Commun. 2014;455:43–57.
26. Lee IS, Lee H, Hur H, Kanda M, Yook JH, Kim BS, et al. Transcriptomic profiling identifies a risk stratification signature for predicting peritoneal recurrence and micrometastasis in gastric cancer. Clin Cancer Res. 2021;27:2292–300.
27. Lee IS, Sahu D, Hur H, Yook JH, Kim BS, Goel A. Discovery and validation of an expression signature for recurrence prediction in high-risk diffuse-type gastric cancer. Gastric Cancer. 2021;24:655–65.
28. Woo Y, Behrendt CE, Yang A, Hahn M, Goel A, Li H, et al. Tumor epigenetic signature and survival in resected gastric cancer patients. J Am Coll Surg. 2021;232:483.e1–91.e1.
29. Roy S, Kanda M, Nomura S, Zhu Z, Toiyama Y, Taketomi A, et al. Diagnostic efficacy of circular RNAs as noninvasive, liquid biopsy biomarkers for early detection of gastric cancer. Mol Cancer. 2022;21:42.
30. Lee IS, Zhu Z, Lee J, Park JO, Wu X, Ong T, et al. A liquid biopsy signature predicts treatment response to fluoropyrimidine plus platinum therapy in patients with metastatic or unresectable gastric cancer: implications for precision oncology. Mol Cancer. 2022;21:9.
31. Shimura T, Kandimalla R, Okugawa Y, Ohi M, Toiyama Y, He C, et al. Novel evidence for m(6)A methylation regulators as prognostic biomarkers and FTO as a potential therapeutic target in gastric cancer. Br J Cancer. 2022;126:228–37.
32. Aurello P, Petrucciani N, Antolino L, Giulitti D, D'Angelo F, Ramacciato G. Follow-up after curative resection for gastric cancer: Is it time to tailor it? World J Gastroenterol. 2017;23:3379–87.
33. Diniz TP, da Costa WL Jr, Gomes CC, de Jesus VHF, Felismino TC, Torres SM, et al. Symptomatic recurrence and survival outcomes after curative treatment of gastric cancer: does intensive follow-up evaluation improve survival? Ann Surg Oncol. 2022;29:274–84.
34. Lu TX, Rothenberg ME. MicroRNA. J Allergy Clin Immunol. 2018;141:1202–7.
35. Peng Y, Croce CM. The role of MicroRNAs in human cancer. Signal Transduct Target Ther. 2016;1:15004.
36. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, et al. Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci USA. 2008;105:10513–8.
37. Shigeyasu K, Nagasaka T, Mori Y, Yokomichi N, Kawai T, Fuji T, et al. Clinical significance of MLH1 methylation and CpG island methylator phenotype as prognostic markers in patients with gastric cancer. PLoS ONE. 2015;10:e0130409.
38. Shigeyasu K, Toden S, Zumwalt TJ, Okugawa Y, Goel A. Emerging role of microRNAs as liquid biopsy biomarkers in gastrointestinal cancers. Clin Cancer Res. 2017;23:2391–9.
39. Jung G, Hernandez-Illan E, Moreira L, Balaguer F, Goel A. Epigenetics of colorectal cancer: biomarker and therapeutic potential. Nat Rev Gastroenterol Hepatol. 2020;17:111–30.
40. Kim JW, Cho SY, Chae J, Kim JW, Kim TY, Lee KW, et al. Adjuvant chemotherapy in microsatellite instability-high gastric cancer. Cancer Res Treat. 2020;52:1178–87.
41. Pietrantonio F, Miceli R, Raimondi A, Kim YW, Kang WK, Langley RE, et al. Individual patient data meta-analysis of the value of microsatellite instability as a biomarker in gastric cancer. J Clin Oncol. 2019;37:3392–400.
42. Cristescu R, Lee J, Nebozhyn M, Kim KM, Ting JC, Wong SS, et al. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. Nat Med. 2015;21:449–56.
43. Kuhn M. Building predictive models in R using the caret package. J Stat Softw. 2008;28:1–26.
44. Mevik BH, Wehrens R. The pls package: principal component and partial least squares regression in R. J Stat Softw. 2007;18:1–23.
45. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw. 2010;33:1–22.
46. Japanese gastric cancer treatment guidelines 2014 (ver. 4). Gastric Cancer. 2017;20:1–19.
47. Sakuramoto S, Sasako M, Yamaguchi T, Kinoshita T, Fujii M, Nashimoto A, et al. Adjuvant chemotherapy for gastric cancer with S-1, an oral fluoropyrimidine. N Engl J Med. 2007;357:1810–20.
48. Nishiwada S, Sho M, Banwait JK, Yamamura K, Akahori T, Nakamura K, et al. A microRNA signature identifies pancreatic ductal adenocarcinoma patients at risk for lymph node metastases. Gastroenterology 2020;159:562–74.
49. Goel A, Nagasaka T, Hamelin R, Boland CR. An optimized pentaplex PCR for detecting DNA mismatch repair-deficient colorectal cancers. PLoS ONE. 2010;5:e9393.
50. Ashktorab H, Smoot DT, Farzanmehr H, Fidelia-Lambert M, Momen B, Hylind L, et al. Clinicopathological features and microsatellite instability (MSI) in colorectal cancers from African Americans. Int J Cancer. 2005;116:914–9.
51. Kanda Y. Investigation of the freely available easy-to-use software 'EZR' for medical statistics. Bone Marrow Transplant. 2013;48:452–8.
52. Youden WJ. Index for rating diagnostic tests. Cancer. 1950;3:32–5.
53. D'Angelica M, Gonen M, Brennan MF, Turnbull AD, Bains M, Karpeh MS. Patterns of initial recurrence in completely resected gastric adenocarcinoma. Ann Surg. 2004;240:808–16.
54. Ito S, Ohashi Y, Sasako M. Survival after recurrence in patients with gastric cancer who receive S-1 adjuvant chemotherapy: exploratory analysis of the ACTS-GC trial. BMC Cancer. 2018;18:449.
55. Chen Y, Wang X. miRDB: an online database for prediction of functional microRNA targets. Nucleic Acids Res. 2020;48:D127–31.
56. Liu W, Wang X. Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. Genome Biol. 2019;20:18.
57. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res. 2022;50:W216–21.
58. Al-Batran SE, Homann N, Pauligk C, Goetze TO, Meiler J, Kasper S, et al. Perioperative chemotherapy with fluorouracil plus leucovorin, oxaliplatin, and docetaxel versus fluorouracil or capecitabine plus cisplatin and epirubicin for locally advanced, resectable gastric or gastro-oesophageal junction adenocarcinoma (FLOT4): a randomised, phase 2/3 trial. Lancet. 2019;393:1948–57.
59. Ychou M, Boige V, Pignon JP, Conroy T, Bouché O, Lebreton G, et al. Perioperative chemotherapy compared with surgery alone for resectable gastroesophageal adenocarcinoma: an FNCLCC and FFCD multicenter phase III trial. J Clin Oncol. 2011;29:1715–21.

ACKNOWLEDGEMENTS

The authors thank Dr. Kota Nakamura, Dr. Katsuki Miyazaki, Dr. Silei Sui, Dr. Caiming Xu, Dr. Hiroyuki Hanayama, Dr. Milad Moloudizargari and Dr. Nour-Lynn Mouallem for their thoughtful discussions and advice during the course of this project.

AUTHOR CONTRIBUTIONS

Study concept and design (KO, SW, MT, Y Kinugasa, AG); provision of samples (KO, MK, MT, Y Kodera, Y Kinugasa); acquisition of clinical data (KO, MK, MT, Y Kodera, Y Kinugasa, AG); analysis and interpretation of data (KO, SW, SR, AG); statistical analysis (KO, SW, SR, AG); drafting of the manuscript (KO, SW, SR, MK, MT, Y Kodera, Y Kinugasa, AG).

FUNDING

This work was supported by the CA072851, CA181572, and CA187956 grants from the National Cancer Institute, National Institutes of Health, and a generous pilot grant from the Stupid Strong Foundation.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

A written informed consent was obtained from all patients who enrolled at the present study. The study was conducted in compliance with the Declaration of Helsinki and was approved by the Institutional Review Boards of all participating institutions.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41416-022-02138-1.

Correspondence and requests for materials should be addressed to Ajay Goel.