

# Gene expression profiles for a prognostic immunoscore in gastric cancer

D. Zeng<sup>1</sup>, R. Zhou<sup>1</sup>, Y. Yu<sup>3</sup>, Y. Luo<sup>1</sup>, J. Zhang<sup>1</sup>, H. Sun<sup>1</sup>, J. Bin<sup>2</sup>, Y. Liao<sup>2</sup>, J. Rao<sup>4</sup>, Y. Zhang<sup>5</sup> and W. Liao<sup>1</sup>

Departments of <sup>1</sup>Oncology and <sup>2</sup>Cardiology, Nanfang Hospital, Southern Medical University, <sup>3</sup>Guangdong Provincial Key Laboratory of Malignant Tumour Epigenetics and Gene Regulation, Breast Tumour Centre, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, and <sup>4</sup>Key Laboratory of New Drug Screening of Guangdong Province, School of Pharmaceutical Sciences, and <sup>5</sup>Key Laboratory of Zebrafish Modelling and Drug Screening for Human Diseases of Guangdong Higher Education Institutes, Department of Developmental Biology, School of Basic Medical Sciences, Southern Medical University, Guangzhou, Guangdong, China

Correspondence to: Professor W. Liao, Department of Oncology, Nanfang Hospital, Southern Medical University, 1838 North Guangzhou Avenue, Guangzhou 510515, China (e-mail: nfyyliaowj@163.com)

**Background:** Increasing evidence has indicated an association between immune infiltration in gastric cancer and clinical outcome. However, reliable prognostic signatures, based on systematic assessments of the immune landscape inferred from bulk tumour transcriptomes, have not been established. The aim was to develop an immune signature, based on the cellular composition of the immune infiltrate inferred from bulk tumour transcriptomes, to improve the prognostic predictions of gastric cancer.

**Methods:** Twenty-two types of immune cell fraction were estimated based on large public gastric cancer cohorts from the Gene Expression Omnibus using CIBERSORT. An immunoscore based on the fraction of immune cell types was then constructed using a least absolute shrinkage and selection operator (LASSO) Cox regression model.

**Results:** Using the LASSO model, an immunoscore was established consisting of 11 types of immune cell fraction. In the training cohort (490 patients), significant differences were found between high- and low-immunoscore groups in overall survival across and within subpopulations with an identical TNM stage. Multivariable analysis revealed that the immunoscore was an independent prognostic factor (hazard ratio 1.92, 95 per cent c.i. 1.54 to 2.40). The prognostic value of the immunoscore was also confirmed in the validation (210) and entire (700) cohorts.

**Conclusion:** The proposed immunoscore represents a promising signature for estimating overall survival in patients with gastric cancer.

## Surgical relevance

Immune infiltration in gastric cancer tissue may predict clinical outcomes and may be used as a prognostic marker.

An immunoscore was constructed based on systematic assessments of the immune landscape inferred from computational analysis of gene expression profiles. The proposed immunoscore

was an independent adverse prognostic factor for overall survival. Patients with stage II and III gastric cancer and a low immunoscore exhibited a more favourable response to adjuvant chemotherapy.

The immunoscore may serve as a biomarker for prognosis and therapeutic outcome in patients with gastric cancer.

Paper accepted 7 March 2018

Published online 25 April 2018 in Wiley Online Library (www.bjs.co.uk). DOI: 10.1002/bjs.10871

## Introduction

Wide variation in clinical outcomes has been reported among patients with gastric cancer who had the same TNM stage and received similar treatment regimens<sup>1,2</sup>, highlighting that TNM staging alone provides incomplete

clinical information. As increasing evidence has suggested the clinical importance of immune infiltration in gastric cancer tissues<sup>3–6</sup>, incorporating the survival impact of immune cells into the TNM staging system may help clinicians predict patient outcomes more reliably and precisely.

The immune response is characterized by numerous types of cell, such as cytotoxic lymphocytes, myeloid cells and antigen-presenting cells, which interact in a highly coordinated manner. However, their prognostic impact differs depending on the type of cancer and the stage<sup>7</sup>. Therefore, enumerating the immune components according to their individual and specialized functions using computer-based analysis may be essential for improving studies of the diverse immune response in gastric cancer and management of its future clinical implementation.

CIBERSORT is a newly proposed computational algorithm for enumeration of immune cell subsets using RNA specimens from multiple tissue types, including solid tumours, and has outperformed other methods regarding noise, unknown mixture content and closely related cell types<sup>8</sup>. In the present study, CIBERSORT was used to estimate the fractions of 22 immune cell types based on clinically annotated gastric cancer gene expression profiling series. Least absolute shrinkage and selection operator (LASSO) Cox regression analysis was used to establish an immunoscore, to provide a statistically powerful means of predicting survival of patients with gastric cancer.

## Methods

### Search and collection of gastric cancer gene expression series

To identify gastric cancer gene expression data with relevant clinicopathological data, systematic computerized searches of Gene Expression Omnibus (GEO) data sets (<https://www.ncbi.nlm.nih.gov/geo/>) were conducted. The search strategy used for identifying eligible series and search results specifically is provided in *Appendix S1* (supporting information). Inclusion and exclusion criteria at each stage of series collection are shown in *Fig. 1*. All candidate series were assessed by two independent reviewers. These series were checked independently for inclusion criteria. Any disagreements were resolved by consensus with a third reviewer.

### Collection of clinical data

The relevant clinical data from these series were retrieved and organized manually when available. For some series, clinical data that were not attached to gene expression profiles were obtained in three ways: downloaded directly from the relevant item page in the GEO data set website; from supplementary material in the relevant literature; and using the GEOquery package of R software (R Project for Statistical Computing, Vienna, Austria). Corresponding authors were contacted for further information where necessary.

### Microarray data processing

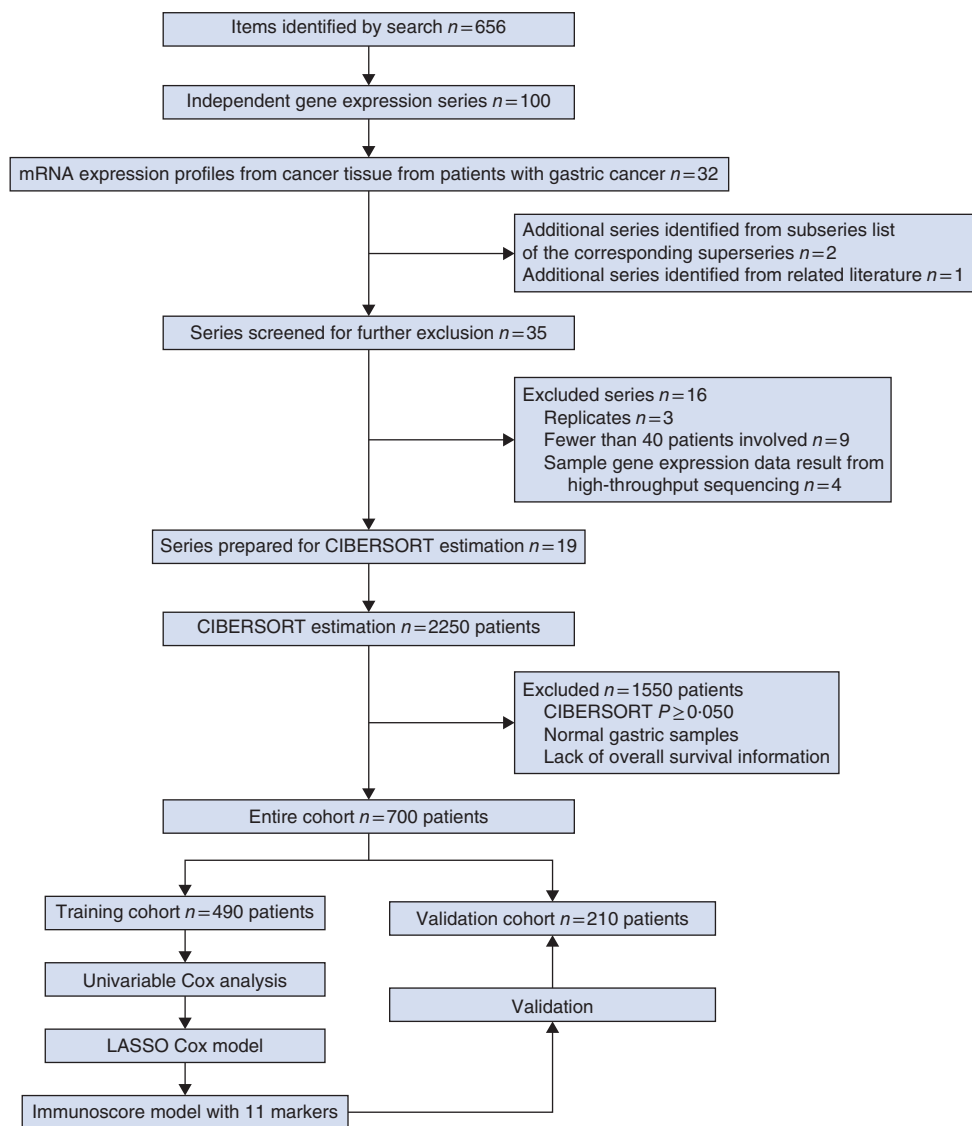
Raw microarray data from Affymetrix<sup>®</sup> (Affymetrix, Santa Clara, California, USA) were downloaded and normalized using a robust multiarray averaging method<sup>9</sup>. The *affy* and *simpleaffy* packages were applied for normalization of Affymetrix data. For gene expression profiles of platforms other than Affymetrix, normalized matrix files were downloaded directly.

### Estimation of immune cell type fractions

To quantify the proportions of immune cells in the gastric cancer samples, the CIBERSORT method and the LM22 gene signature were used<sup>8</sup>; the latter contains 547 genes and allows highly sensitive and specific discrimination of 22 human haematopoietic cell phenotypes including B cells, T cells, natural killer cells, macrophages, dendritic cells and myeloid subsets. The CIBERSORT method is well designed and has been validated on gene expression profiles measured using microarrays. CIBERSORT derives a *P* value for the deconvolution of each sample using Monte Carlo sampling, providing a measure of confidence in the results. At a threshold of  $P < 0.050$ , the results of the inferred fractions of immune cell populations produced by CIBERSORT were considered accurate<sup>10</sup>. Based on this observation, only patients with a CIBERSORT  $P < 0.050$  were considered eligible for further analysis. The proportions of immune cells were predicted separately for each gene expression series. For each sample, the sum of all estimates of immune cell type fractions equalled 1.

### Study population and clinicopathological variables

Patients with CIBERSORT  $P \geq 0.050$  were excluded, as were those with normal gastric samples and patients for whom survival information was lacking. Clinicopathological information was collected including: patient age, sex, TNM stage, tumour grade, Laurén classification, primary tumour site, whether adjuvant chemotherapy was administered and regimen used, survival duration in months, and survival status at date of last follow-up. Data on the Asian Cancer Research Group (ACRG) molecular subtypes<sup>11</sup>, including microsatellite instability (MSI), epithelial-to-mesenchymal transition (EMT), microsatellite stable (MSS)/TP53– and MSS/TP53+, were also retrieved where available. Among these factors, sex, Laurén classification, tumour site, tumour grade, TNM stage, treatment type and ACRG molecular subtypes were considered as categorical variables. Age was considered as a continuous variable. The seventh edition of the TNM staging system<sup>12</sup> was used only for patients in the GSE29272 series, and no specific T, N and M



**Fig. 1** Flow chart of data collection and analysis. LASSO, least absolute shrinkage and selection operator

categories were provided for these patients. For all other patients, tumours were staged according to the sixth edition of the TNM staging system<sup>13</sup>. None of the analyses and discussion related to TNM stage in this study included patients from the GSE29272 series.

### Random grouping method

The patients were separated into training and validation sets in a ratio of 7:3 using the stratified randomization method. This involved generating random values from a normal distribution with specified mean (0) and standard deviation (1) values in each gene expression series included in model construction, and ordering them from high to low. The top 70 per cent of patients in each gene expression

series was included in the training cohort used to identify and evaluate predictors, and the remaining 30 per cent as the validation cohort used to validate the final model.

### Primary outcome

Information on overall survival, defined as the interval between date of diagnosis and date of death from any cause, was documented for the majority of patients and used as the primary endpoint.

### Statistical analysis

Group comparisons were performed for continuous and categorical variables using one-way ANOVA and the  $\chi^2$

test respectively. Correlations between the immunoscore and mRNA expression of genes were analysed by means of Pearson's correlation test. Survival curves were constructed by the Kaplan–Meier method and compared by means of the log rank test. Hazard ratios for univariable analyses were calculated using a univariable Cox proportional hazards regression model. The penalized Cox regression model with LASSO penalty was used to select the most useful prognostic markers among 22 immune cell subsets, and the optimal values of the penalty parameter  $\lambda$  were determined by tenfold cross-validations<sup>14</sup>. An immunoscore model was then constructed based on the fraction of the selected immune cells using Cox regression coefficients in the training cohort. Of note, immune cell fractions were all analysed as binary variables in LASSO; the optimal cut-off values were evaluated based on the association between overall survival and cell fraction in the training cohort using the *survminer* package. A multivariable Cox regression model with the enter method was used to determine independent prognostic factors. Only patients with complete clinical information were included in multivariable survival analyses; those with any missing value were excluded. The sensitivity and specificity of the survival prediction based on the immunoscore were depicted by a time-dependent receiver operating characteristic (ROC) curve, with quantification of the area under the ROC curve using the *timeROC* package<sup>15</sup>. The discrimination of the prognostic models was measured and compared by means of Harrell's concordance index (C-index), using the *survival* package in cohorts of patients with and without stage IV disease. Details of methods for nomogram construction and validation are provided in *Appendix S1* (supporting information). All statistical tests were two-sided and  $P < 0.050$  was considered statistically significant. Statistical analyses were conducted using R software and SPSS<sup>®</sup> version 19.0 (IBM, Armonk, New York, USA). This study was conducted and reported in line with the TRIPOD guidelines<sup>16</sup>.

## Results

The patient selection scheme is shown in *Fig. 1*. After applying data filter criteria, 700 clinically annotated gastric cancer samples with overall survival information were available for further analyses. Patient characteristics are detailed in *Table 1*. Data from 382 patients (54.6 per cent) were censored. *Fig. S1* (supporting information) provides a summary of the immune cell composition within and across clinical subgroups of gastric cancer tissues in the entire cohort. In general, the five most common immune cell fractions in gastric cancer tissues were plasma cells, M2 macrophages, M1 macrophages, resting memory CD4+ T cells and CD8+ T cells, and the sum of their

**Table 1** Baseline patient characteristics

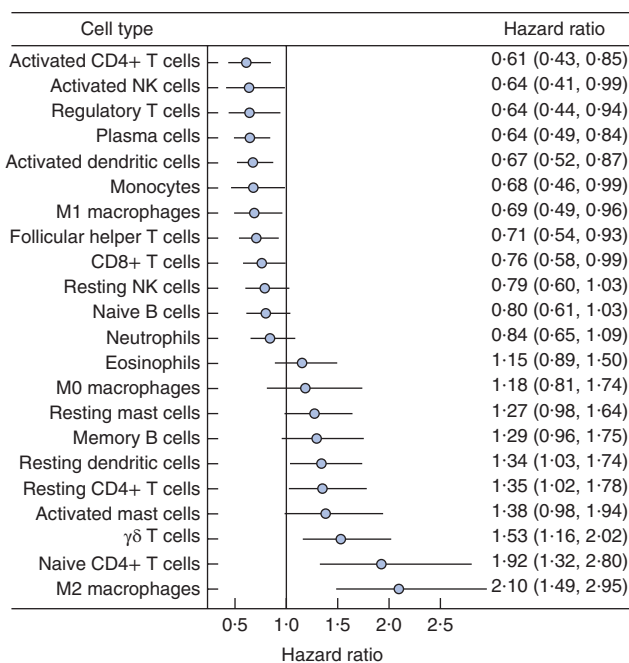
	No. of patients (n = 700)
Affymetrix <sup>®</sup> platform	
HG-U133 Plus 2.0	566 (80.9)
HG-U133A	134 (19.1)
Age (years)	
18–64	370 (52.9)
> 64	330 (47.1)
Sex ratio (M : F)	478 : 222
Tumour stage*	
I	66 (9.4)
II	140 (20.0)
III	211 (30.1)
IV	157 (22.4)
Unknown	126 (18.0)
Tumour grade	
Well/moderately differentiated	140 (20.0)
Poorly differentiated	164 (23.4)
Unknown	396 (56.6)
Laurén classification	
Intestinal	331 (47.3)
Diffuse	241 (34.4)
Mixed	57 (8.1)
Unknown	71 (10.1)
Tumour site	
Cardia	85 (12.1)
Body	113 (16.1)
Antrum	150 (21.4)
Unknown	352 (50.3)
Adjuvant chemotherapy	
Yes	162 (23.1)
No	137 (19.6)
Unknown	401 (57.3)

Values in parentheses are percentages. \*TNM sixth edition.

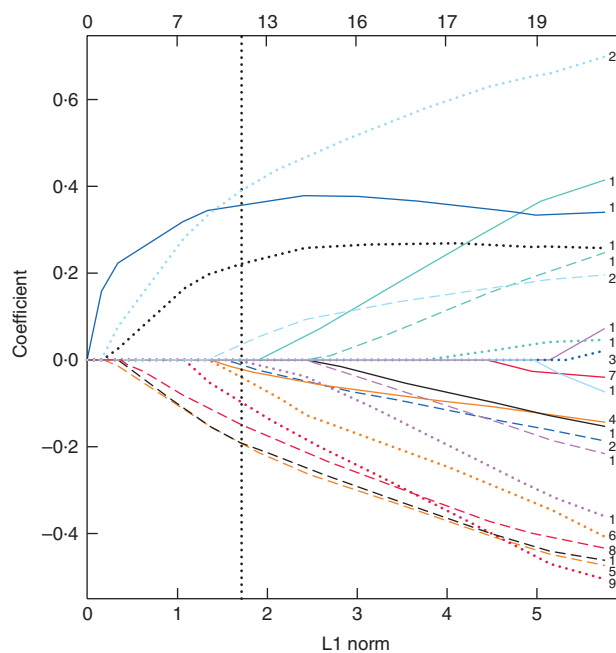
mean proportions was more than 50 per cent. Such cell composition patterns were observed in all clinical subgroups.

## Derivation of the immunoscore

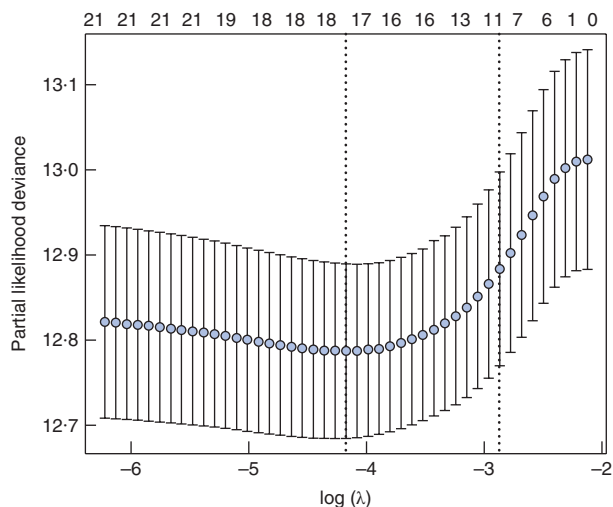
The *survminer* package was used to generate the optimal cut-off values for the fraction of each immune cell in the training cohort (490 patients) (*Table S1*, supporting information). *Fig. 2a* shows a forest plot of the associations between each of the immune cell subsets and overall survival. LASSO Cox regression analysis was used to build an immunoscore model in the training cohort (*Fig. 2b,c*). The formula for the immunoscore can be found in *Appendix S1* (supporting information). In this formula, the immune cell fraction level was valued as 0 or 1; a value of 0 was assigned when the fraction of one type of cell was less than the corresponding cut-off value, and a value of 1 otherwise. The prognostic accuracy of the immunoscore, assessed as a continuous variable, was investigated in the training cohort by using time-dependent ROC analysis at the time points 2, 3 and 5 years (*Fig. 2d*). Patients



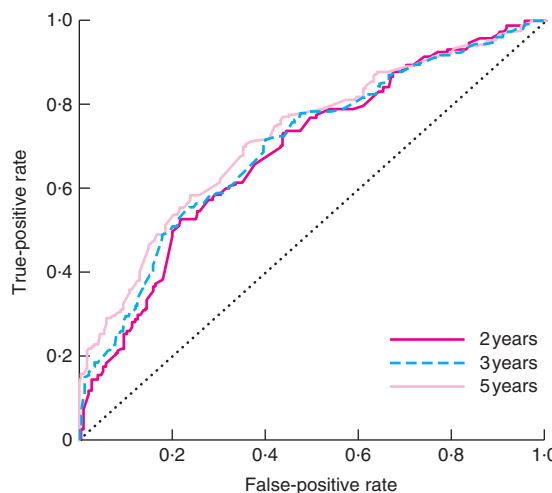
**a** Forest plots



**b** LASSO coefficient profile

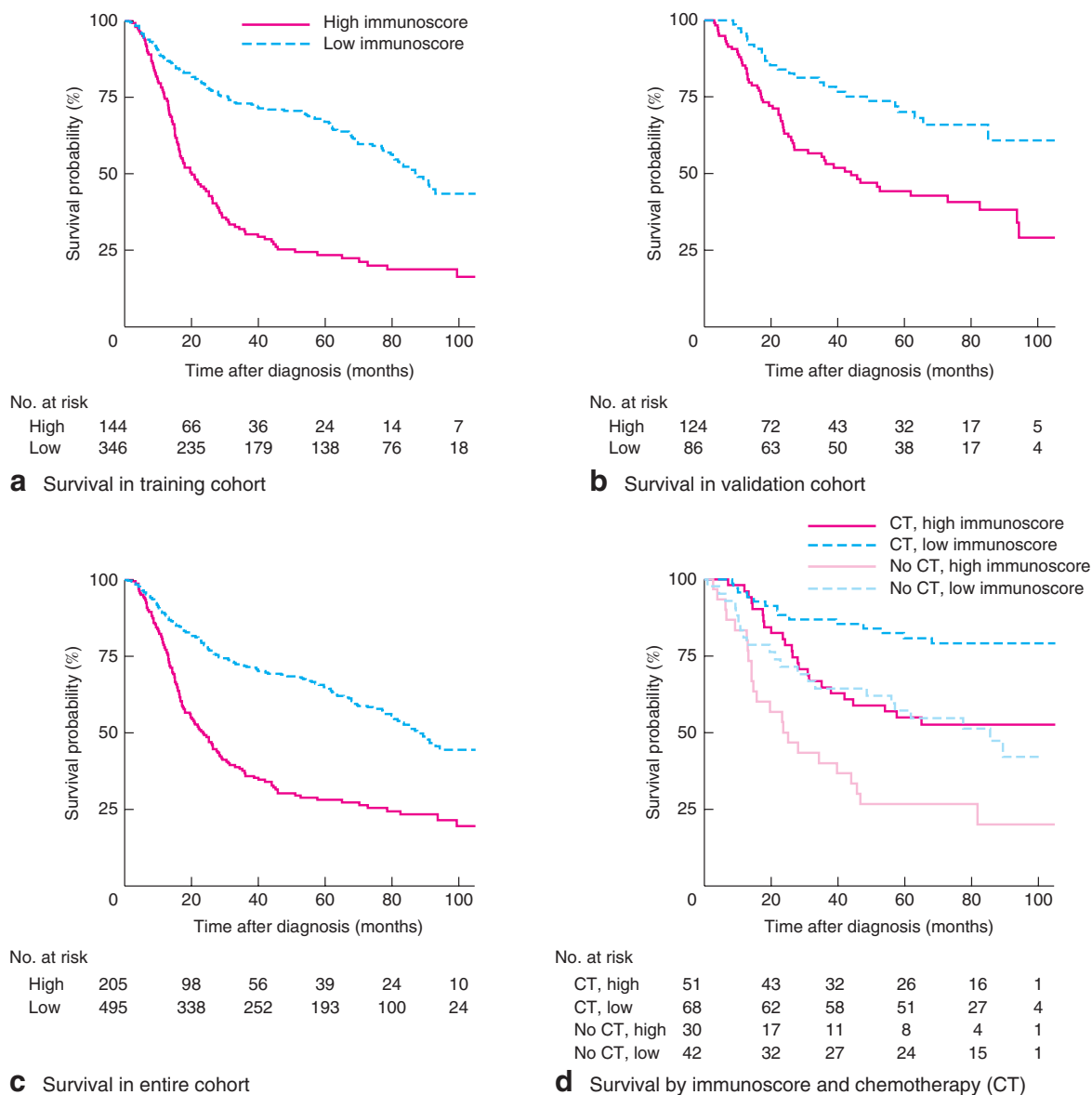


**c** Cross-validation



**d** Time-dependent ROC curves

**Fig. 2** Construction of the immunoscore model. **a** Forest plots showing associations between different immune cell subsets and overall survival in the training cohort. Unadjusted hazard ratios are shown with 95 per cent confidence intervals. NK, natural killer. **b** Least absolute shrinkage and selection operator (LASSO) coefficient profiles of the fractions of 21 immune cell types. The dotted line indicates the value chosen by tenfold cross-validation. Immune cell type: 1, M2 macrophages; 2, M1 macrophages; 3, M0 macrophages; 4, CD8+ T cells; 5, activated memory CD4+ T cells; 6, regulatory T cells; 7, neutrophils; 8, activated dendritic cells; 9, monocytes; 10, follicular helper T cells; 11, resting NK cells; 12, activated NK cells; 13, activated mast cells; 14, resting mast cells; 15, memory B cells; 16, naive B cells; 17, plasma cells; 18,  $\gamma\delta$  T cells; 19, eosinophils; 20, resting dendritic cells; 21, naive CD4+ T cells. **c** Tenfold cross-validation for tuning parameter selection in the LASSO model. The partial likelihood deviance values are shown, with error bars representing s.e. The dotted vertical lines are drawn at the optimal values by minimum criteria and 1 – s.e. criteria. In **b** and **c**, the numbers above the graph represent the number of cell types involved in the LASSO model. **d** Immunoscore measured by time-dependent receiver–operating characteristic (ROC) curves in the training cohort. The area under the ROC curve was 0.68, 0.69 and 0.72 for the immunoscore at 2, 3 and 5 years respectively



**Fig. 3** Survival impact of the immunoscore. **a–d** Kaplan–Meier curves for overall survival by immunoscore group in the training cohort (**a**), validation cohort (**b**) and entire cohort (**c**), and for patients with stage II–III gastric cancer in subgroups stratified by both receipt of adjuvant chemotherapy (CT) and immunoscore (**d**). Hazard ratios are shown with 95 per cent confidence intervals.  $P < 0.001$  (log rank test)

in the training cohort were then assigned to a high- or low-immunoscore group using the cut-off value ( $-0.37$ ) obtained with the *survminer* package. Five-year survival rates were 67.0 and 23.6 per cent respectively for the low- and high-immunoscore groups (hazard ratio (HR) 2.93, 95 per cent c.i. 2.26 to 3.80) (*Fig. 3a*). The association between the immunoscore and overall survival was also significant when evaluated as a continuous variable in the multivariable Cox regression model (HR 1.92, 1.54 to 2.40) (*Table 2*).

The results of univariable analyses of clinicopathological variables are shown in *Table S2* (supporting information).

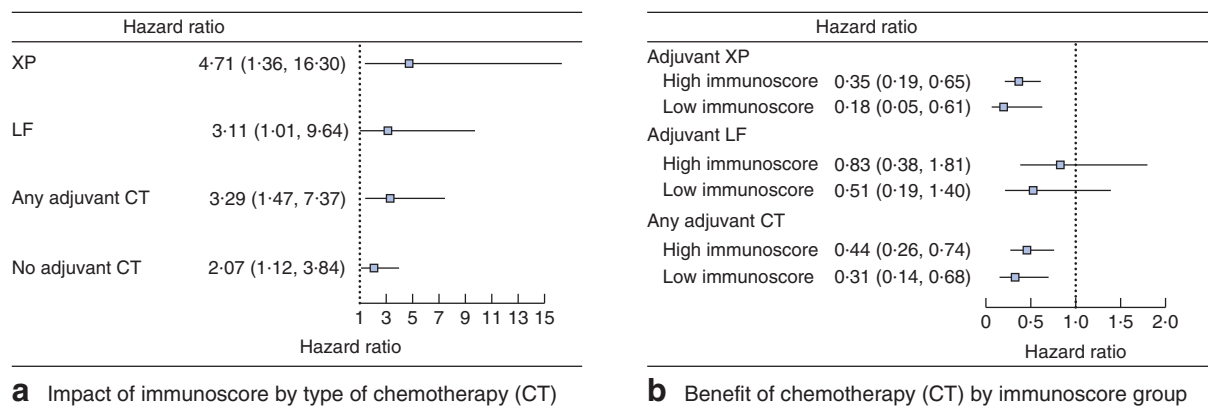
### Validation of immunoscore for predicting survival in the validation cohort and entire cohort

To confirm that the proposed immunoscore model has similar prognostic value in different populations, the same formula was applied to the validation cohort and

**Table 2** Results of multivariable Cox regression analysis

	Training cohort		Validation cohort		Entire cohort	
	Hazard ratio	P	Hazard ratio	P	Hazard ratio	P
Immunoscore*	1.92 (1.54, 2.40)	< 0.001	1.81 (1.17, 2.79)	0.008	1.89 (1.56, 2.30)	< 0.001
Age*	1.02 (1.01, 1.04)	< 0.001	1.02 (1.00, 1.04)	0.069	1.02 (1.01, 1.03)	< 0.001
Tumour stage		< 0.001		< 0.001		< 0.001
I	1.00 (reference)		1.00 (reference)		1.00 (reference)	
II	3.23 (1.24, 8.41)	0.016	2.12 (0.71, 6.34)	0.179	2.75 (1.35, 5.64)	0.006
III	7.85 (3.14, 19.62)	< 0.001	2.59 (0.88, 7.59)	0.083	5.36 (2.69, 10.67)	< 0.001
IV	22.21 (8.85, 55.73)	< 0.001	6.58 (2.25, 19.19)	0.001	14.34 (7.18, 28.64)	< 0.001
Laurén classification		0.391		0.670		0.536
Intestinal	1.00 (reference)		1.00 (reference)		1.00 (reference)	
Diffuse	1.41 (0.86, 2.33)	0.177	1.09 (0.52, 2.25)	0.825	1.18 (0.78, 1.76)	0.433
Mixed	1.04 (0.75, 1.42)	0.825	1.26 (0.76, 2.10)	0.374	1.11 (0.85, 1.44)	0.458

Values in parentheses are 95 per cent confidence intervals. \*Continuous variable.



**Fig. 4** **a** Forest plot showing the survival impact of immunoscore stratified by type of adjuvant chemotherapy (CT) among patients with stage II and III gastric cancer. Hazard ratios, with 95 per cent confidence intervals, are shown for the high immunoscore group *versus* the low immunoscore group. **b** Forest plot showing the benefit of CT in different immunoscore groups of patients with stage II and III gastric cancer. Hazard ratios, with 95 per cent confidence intervals, are shown for CT *versus* no CT in each immunoscore group. XP, xeloda plus cisplatin; LF, leucovorin plus fluorouracil

also to the entire cohort. The prognostic accuracy of the immunoscore as a continuous variable in these cohorts was also assessed using time-dependent ROC analysis (Fig. S2, supporting information). The patients were assigned to a high- or low-immunoscore group using the cut-off value obtained from the corresponding cohort (validation,  $-0.82$ ; entire,  $-0.37$ ). Consistent with the findings in the training cohort, patients in the high-immunoscore group had a significantly lower overall survival rate than those in the low-immunoscore group in both the validation cohort (HR 2.38, 95 per cent c.i. 1.46 to 3.77) (Fig. 3b) and the entire cohort (HR 2.54, 2.03 to 3.17) (Fig. 3c). The immunoscore model was also demonstrated to be an independent prognostic factor when analysed as a continuous variable in multivariable analysis using both the validation cohort (HR 1.81, 1.17 to 2.79) and the entire cohort (HR 1.89, 1.56 to 2.30) (Table 2).

### Immunoscore and TNM staging

Stratification analyses were performed in the entire cohort of patients grouped by TNM stage. The immunoscore as a continuous variable identified patients with different prognoses in each TNM stage subgroup, although the result was not significant for stage I ( $P = 0.072$ ) (Fig. S3, supporting information). Based on comparisons between the immunoscore and TNM stage in the training cohort, it was found that the ability of the immunoscore to predict survival was inferior to that of the TNM stage for patients with stage I–IV gastric cancer, whereas the two were similar for patients with stage I–III disease (Table S3, supporting information). A similar tendency was also observed in both the validation and entire cohorts (Table S3, supporting information).

A nomogram that integrated both the immunoscore and TNM stage showed improved prognostic accuracy in the

training, validation and entire cohorts compared with that of TNM stage alone (*Appendix S1, Fig. S4, Fig. S5 and Table S4*, supporting information).

### Immunoscore and adjuvant chemotherapy

Information regarding the administration of adjuvant chemotherapy was documented only in the GSE62254 series (299 patients). The immunoscore was therefore used specifically on patients with stage II–III disease in this series to explore whether application of adjuvant chemotherapy and differences in chemotherapy regimen would affect the ability of the immunoscore to predict survival. Patients were assigned to high- and low-immunoscore groups, with an immunoscore of  $-0.82$  as the cut-off. The survival advantage for the low-immunoscore group was evident both in patients who received chemotherapy and those who did not, regardless of the chemotherapy regimen (*Fig. 3d, Fig. 4a*). Moreover, the significant benefit of adjuvant chemotherapy for survival was observed in the high- and low-immunoscore groups, and was more obvious for patients with a low immunoscore. This tendency was noted, regardless of whether the patient received a xeloda plus cisplatin or leucovorin plus fluorouracil regimen (*Fig. 4b*). After multivariable adjustment for clinicopathological variables, molecular subtypes and adjuvant chemotherapy status, the immunoscore remained a powerful and independent predictor (HR 2.50, 95 per cent c.i. 1.57 to 3.98) (*Table S5*, supporting information). The corresponding nomogram confirmed that the immunoscore contributed much more to prognosis than adjuvant chemotherapy status and TNM stage (*Appendix S1, Fig. S6 and Table S6*, supporting information).

### Distribution of immunoscore, clinical characteristics and molecular subtypes

The distribution of the immunoscore, clinical characteristics and molecular subtypes was assessed in the GSE62254 series, which contained the most comprehensive patient information (*Fig. S7, Table S7 and Table S8*, supporting information). In terms of clinical characteristics, a higher immunoscore was significantly associated with poorer tumour differentiation, more advanced TNM stage and death (*Fig. S7a and Table S7*, supporting information). The immunoscore in patients with diffuse-type gastric cancer was significantly higher than that in patients with the intestinal and mixed types. However, the distribution of the immunoscore value did not vary significantly among subgroups based on tumour site, *Helicobacter pylori* infection status and Epstein–Barr virus infection status.

Regarding associations between the immunoscore and ACRG molecular subtypes, the immunoscore value was highest in patients with an EMT subtype, and lowest in those with MSI; there was no significant difference in the immunoscore value between MSS/TP53– and MSS/TP53+ subtypes. Notably, there was a significant negative correlation between the immunoscore value and gene expression of certain immune checkpoint regulators and inflammatory mediators, including *PD-L1* ( $P=0.014$ ), *CD47* ( $P<0.001$ ), *CLTA4* ( $P=0.015$ ), *LAG3* ( $P<0.001$ ), *IFNG* ( $P<0.001$ ), *GZMB* ( $P<0.001$ ), *TNFA* ( $P=0.009$ ), *IL-1B* ( $P=0.010$ ) and *IL-12A* ( $P=0.019$ ), whereas *IL-1A*, *IL-6* and *IL-12B* showed no significant correlation (*Table S8*, supporting information). The correlation between immunoscore value and all genes that the Affymetrix® microarray detected was also analysed; an electronic link to the results, showing the genes with a statistically significant correlation, is available in *Appendix S1* (supporting information).

### Discussion

The immunoscore, a novel prognostic tool designed to improve survival prediction after diagnosis of gastric cancer, was developed and validated in this retrospective study. The immunoscore is based on the fractions of 11 immune cells. The results showed a clear separation of overall survival curves between patients with high and low immunoscores. Furthermore, the immunoscore predicted survival in groups of patients with identical TNM stage, suggesting that this model could have prognostic value that complements TNM staging.

In recent years, immune profiling studies have gained a forefront position in cancer research<sup>7,17</sup>. Several models<sup>17–19</sup> based on immunoscore have been reported to quantify the immune contexture and to provide a statistically strong parameter for prognosis in patients with various types of solid tumour, including gastric cancer<sup>2</sup>. In these studies, immunohistochemistry was one common research strategy for studying cell heterogeneity. However, immunohistochemistry relies heavily on a limited repertoire of phenotypic markers and biopsy specimens of sufficient size. Owing to technical restrictions, these studies were always limited by either small sample size, few cell types, or both. Moreover, a standardized and reproducible measurement of the intensity of staining, and hence quantitation of protein expression, is also intrinsically difficult in immunohistochemistry.

In contrast to previous studies, the candidate immune cells used to build the present immunoscore model were estimated based on a high-throughput gene expression



profile generated using the newly developed algorithm CIBERSORT. By applying this computer-based analytical method to public genomic data downloaded from the GEO data set, it was possible to generate an expanded view of the immune response at cellular level, which allowed precise investigation of more cell subtypes and more specific functional phenotypes within a large patient cohort than achieved in previous studies. With further use of LASSO Cox regression models, as a statistical method for screening cell variables to establish the immunoscore model, the predictive accuracy could be improved significantly<sup>2,20–22</sup>. The C-index suggested that the predictive ability of the immunoscore for survival was inferior to that of TNM stage if the cohort included patients with metastatic disease (stage I–IV), but was similar to that of TNM stage in patients without distant metastasis (stage I–III). This could be attributed to the universally recognized poor prognosis associated with distant metastases and the multiple risk factors affecting prognosis. Therefore, the effect of the immune microenvironment on prognosis for patients with metastatic disease is relatively diminished. However, the nomogram that included both the immunoscore and TNM stage had significantly better prognostic value than TNM stage alone in both stage I–III and stage I–IV cohorts, implying that the immunoscore could be used to reinforce the prognostic ability of TNM staging. The value of the immunoscore was confirmed in the non-overlapping validation cohort and the entire patient cohort, indicating excellent reproducibility for gastric cancer.

Adjuvant chemotherapy is currently regarded as the standard protocol for patients with stage II or III gastric cancer based on National Comprehensive Cancer Network guidelines<sup>23,24</sup>. However, the criteria for selection of candidates who are likely to benefit from adjuvant chemotherapy remain controversial<sup>2,24</sup>. Several studies<sup>2,25,26</sup> have given rise to the hypothesis that tumour infiltration by lymphocytes defines a phenotype of chemotherapy-sensitive disease in multiple types of cancer. Similar to the findings in these reports, the present study found that the reduction in cancer mortality mediated by adjuvant chemotherapy tended to be greater in patients in the low-immunoscore group (more immunity-activating lymphocyte infiltration). One of the underlying mechanisms explaining such a phenomenon is that the interferon secreted by lymphocytes could sensitize cells to chemotherapy<sup>25</sup>. Coincidentally, the present study also showed a significant negative correlation between the immunoscore value and *IFNA2*, *IFNB1* and *IFNG* mRNA expression level. This indicated that interferon secretion might participate in the biological process of chemotherapy sensitization in patients with

gastric cancer and a low immunoscore. Further investigation of the mechanism between the immunoscore and chemosensitivity in gastric cancer may provide additional information and strategies for treatment.

The survival advantage for patients with MSI cancer has been demonstrated in a number of studies<sup>27–29</sup>. Owing to its supposedly high antigenic potential, the MSI subtype is well known to be associated with a high level of lymphocytic infiltration and is a promising candidate for immunotherapy such as anti-PD-1 treatment<sup>30</sup>. It was found that immunoscore values of patients with MSI were significantly lower than those in patients with MSS status. In particular, the immunoscore value was negatively associated with multiple immune checkpoint markers and carcinostatic inflammatory mediators. The authors speculate that immunotherapy might also be a preferable choice for patients in the low-immunoscore group.

This study has some limitations. First, it was based on publicly available data sets, and it was not possible to obtain all information needed for each patient. This suggests the possibility that some patients with acute infection or immune system disorders, or those taking anti-inflammatory drugs, were included in this study; such patients ideally should have been excluded. Second, the methodology for interpreting immune infiltration and the appropriate cut-off value needs to be standardized. Third, given the major clinical importance of distinct tumour regions, it is appropriate to conduct immune infiltration evaluation systematically in the core of the tumour and the invasive margin. However, the gene expression profiles used here were all derived from a core sample of tumour tissue, making it impossible for the location of the immune cell to be taken into account when establishing the immunoscore model. Finally, as all patients in this study were selected retrospectively, the potential bias relating to unbalanced clinicopathological features with treatment heterogeneity cannot be ignored. Further prospective studies are required to validate the results.

### Acknowledgements

D.Z. and R.Z. contributed equally to this study. This work was supported by the National Natural Science Foundation of China (no. 81472314 to W.L.). The authors thank the members of W. Liao's laboratory for advice and discussion; and Editage for language editing.

*Disclosure:* The authors declare no conflict of interest.

### References

- Zhou R, Wu Z, Zhang J, Wang H, Su Y, Huang N *et al.* Clinical significance of accurate identification of lymph node

- status in distant metastatic gastric cancer. *Oncotarget* 2016; **7**: 1029–1041.
- 2 Jiang Y, Zhang Q, Hu Y, Li T, Yu J, Zhao L *et al.* ImmunoScore signature: a prognostic and predictive tool in gastric cancer. *Ann Surg* 2018; **267**: 504–513.
  - 3 Hennequin A, Derangère V, Boidot R, Apetoh L, Vincent J, Orry D *et al.* Tumor infiltration by Tbet+ effector T cells and CD20+ B cells is associated with survival in gastric cancer patients. *Oncoimmunology* 2015; **5**: e1054598.
  - 4 Lee HE, Chae SW, Lee YJ, Kim MA, Lee HS, Lee BL *et al.* Prognostic implications of type and density of tumour-infiltrating lymphocytes in gastric cancer. *Br J Cancer* 2008; **99**: 1704–1711.
  - 5 Zhang R, Li F, Li H, Yu J, Ren X. The clinical significance of memory T cells and its subsets in gastric cancer. *Clin Transl Oncol* 2014; **16**: 257–265.
  - 6 Zhang WJ, Zhou ZH, Guo M, Yang LQ, Xu YY, Pang TH *et al.* High infiltration of polarized CD163+ tumor-associated macrophages correlates with aberrant expressions of CSCs markers, and predicts prognosis in patients with recurrent gastric cancer. *J Cancer* 2017; **8**: 363–370.
  - 7 Fridman WH, Pagès F, Sautès-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer* 2012; **12**: 298–306.
  - 8 Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* 2015; **12**: 453–457.
  - 9 Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003; **4**: 249–264.
  - 10 Ali HR, Chlon L, Pharoah PD, Markowitz F, Caldas C. Patterns of immune infiltration in breast cancer and their clinical implications: a gene-expression-based retrospective study. *PLoS Med* 2016; **13**: e1002194.
  - 11 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**: 449–456.
  - 12 Sobin LH, Gospodarowicz M, Wittekind C. *TNM Classification of Malignant Tumours* (7th edn). Weinheim: Wiley, 2010.
  - 13 Greene FL, Page DL, Fleming ID, Fritz A, Balch CM, Haller DG *et al.* (eds). *AJCC Cancer Staging Handbook: TNM Classification of Malignant Tumors* (6th edn). Springer: New York, 2002.
  - 14 Goeman JJ. L1 penalized estimation in the Cox proportional hazards model. *Biom J* 2010; **52**: 70–84.
  - 15 Kamarudin AN, Cox T, Kolamunnage-Dona R. Time-dependent ROC curve analysis in medical research: current methods and applications. *BMC Med Res Methodol* 2017; **17**: 53.
  - 16 Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Br J Surg* 2015; **102**: 148–158.
  - 17 Galon J, Mlecnik B, Bindea G, Angell HK, Berger A, Lagorce C *et al.* Towards the introduction of the ‘Immunoscore’ in the classification of malignant tumours. *J Pathol* 2014; **232**: 199–209.
  - 18 Angell H, Galon J. From the immune contexture to the Immunoscore: the role of prognostic and predictive immune markers in cancer. *Curr Opin Immunol* 2013; **25**: 261–267.
  - 19 Galon J, Pagès F, Marincola FM, Thurin M, Trinchieri G, Fox BA *et al.* The immune score as a new possible approach for the classification of cancer. *J Transl Med* 2012; **10**: 1.
  - 20 Agesen TH, Sveen A, Merok MA, Lind GE, Nesbakken A, Skotheim RI *et al.* ColoGuideEx: a robust gene classifier specific for stage II colorectal cancer prognosis. *Gut* 2012; **61**: 1560–1567.
  - 21 Zhang JX, Song W, Chen ZH, Wei JH, Liao YJ, Lei J *et al.* Prognostic and predictive value of a microRNA signature in stage II colon cancer: a microRNA expression analysis. *Lancet Oncol* 2013; **14**: 1295–1306.
  - 22 Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med* 1997; **16**: 385–395.
  - 23 GASTRIC (Global Advanced/Adjuvant Stomach Tumor Research International Collaboration) Group, Paoletti X, Oba K, Burzykowski T, Michiels S, Ohashi Y *et al.* Benefit of adjuvant chemotherapy for resectable gastric cancer: a meta-analysis. *JAMA* 2010; **303**: 1729–1737.
  - 24 Nishida T. Adjuvant therapy for gastric cancer after D2 gastrectomy. *Lancet* 2012; **379**: 291–292.
  - 25 Galluzzi L, Buquè A, Kepp O, Zitvogel L, Kroemer G. Immunological effects of conventional chemotherapy and targeted anticancer agents. *Cancer Cell* 2015; **28**: 690–714.
  - 26 Zitvogel L, Apetoh L, Ghiringhelli F, André F, Tesnière A, Kroemer G. The anticancer immune response: indispensable for therapeutic success? *J Clin Invest* 2008; **118**: 1991–2001.
  - 27 An C, Choi IS, Yao JC, Worah S, Xie K, Mansfield PF *et al.* Prognostic significance of CpG island methylator phenotype and microsatellite instability in gastric carcinoma. *Clin Cancer Res* 2005; **11**: 656–663.
  - 28 Giampieri R, Maccaroni E, Mandolesi A, Del Prete M, Andrikou K, Faloppi L *et al.* Mismatch repair deficiency may affect clinical outcome through immune response activation in metastatic gastric cancer patients receiving first-line chemotherapy. *Gastric Cancer* 2017; **20**: 156–163.
  - 29 Kim SY, Choi YY, An JY, Shin HB, Jo A, Choi H *et al.* The benefit of microsatellite instability is attenuated by chemotherapy in stage II and stage III gastric cancer: results from a large cohort with subgroup analyses. *Int J Cancer* 2015; **137**: 819–825.
  - 30 Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK *et al.* Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 2017; **357**: 409–413.

**Supporting information**

Additional supporting information can be found online in the Supporting Information section at the end of the article.