Predictive test for chemotherapy response in resectable gastric cancer: a multi-cohort, retrospective analysis

Jae-Ho Cheong, Han-Kwang Yang, Hyunki Kim, Woo Ho Kim, Young-Woo Kim, Myeong-Cherl Kook, Young-Kyu Park, Hyung-Ho Kim, Hye Seung Lee, Kyung Hee Lee, Mi Jin Gu, Ha Yan Kim, Jinae Lee, Seung Ho Choi, Soonwon Hong, Jong Won Kim, Yoon Young Choi, Woo Jin Hyung, Eunji Jang, Hyeseon Kim, Yong-Min Huh, Sung Hoon Noh**

Summary

Background Adjuvant chemotherapy after surgery improves survival of patients with stage II–III, resectable gastric cancer. However, the overall survival benefit observed after adjuvant chemotherapy is moderate, suggesting that not all patients with resectable gastric cancer treated with adjuvant chemotherapy benefit from it. We aimed to develop and validate a predictive test for adjuvant chemotherapy response in patients with resectable, stage II–III gastric cancer.

Methods In this multi-cohort, retrospective study, we developed through a multi-step strategy a predictive test consisting of two rule-based classifier algorithms with predictive value for adjuvant chemotherapy response and prognosis. Exploratory bioinformatics analyses identified biologically relevant candidate genes in gastric cancer transcriptome datasets. In the discovery analysis, a four-gene, real-time RT-PCR assay was developed and analytically validated in formalin-fixed, paraffin-embedded (FFPE) tumour tissues from an internal cohort of 307 patients with stage II–III gastric cancer treated at the Yonsei Cancer Center with D2 gastrectomy plus adjuvant fluorouracil-based chemotherapy (n=193) or surgery alone (n=114). The same internal cohort was used to evaluate the prognostic and chemotherapy response predictive value of the single patient classifier genes using associations with 5-year overall survival. The results were validated with a subset (n=625) of FFPE tumour samples from an independent cohort of patients treated in the CLASSIC trial (NCT00411229), who received D2 gastrectomy plus capecitabine and oxaliplatin chemotherapy (n=323) or surgery alone (n=302). The primary endpoint was 5-year overall survival.

Findings We identified four classifier genes related to relevant gastric cancer features (*GZMB*, *WARS*, *SFRP4*, and *CDX1*) that formed the single patient classifier assay. In the validation cohort, the prognostic single patient classifier (based on the expression of *GZMB*, *WARS*, and *SFRP4*) identified 79 (13%) of 625 patients as low risk, 296 (47%) as intermediate risk, and 250 (40%) as high risk, and 5-year overall survival for these groups was 83·2% (95% CI 75·2–92·0), 74·8% (69·9–80·1), and 66·0% (60·1–72·4), respectively (p=0·012). The predictive single patient classifier (based on the expression of *GZMB*, *WARS*, and *CDX1*) assigned 281 (45%) of 625 patients in the validation cohort to the chemotherapy-benefit group and 344 (55%) to the no-benefit group. In the predicted chemotherapy-benefit group, 5-year overall survival was significantly improved in those patients who had received adjuvant chemotherapy after surgery compared with those who received surgery only (80% [95% CI 73·5–87·1] *vs* 64·5% [56·8–73·3]; univariate hazard ratio 0·47 [95% CI 0·30–0·75], p=0·0015), whereas no such improvement in 5-year overall survival was observed in the no-benefit group (72·9% [66·5–79·9] in patients who received chemotherapy plus surgery *vs* 72·5% [65·8–79·9] in patients who only had surgery; 0·93 [0·62–1·38], p=0·71). The predictive single patient classifier groups (chemotherapy benefit *vs* no-benefit) could predict adjuvant chemotherapy benefit in terms of 5-year overall survival in the validation cohort ($p_{interaction}$=0·036 in univariate analysis). Similar results were obtained in the internal evaluation cohort.

Interpretation The single patient classifiers validated in this study provide clinically important prognostic information independent of standard risk-stratification methods and predicted chemotherapy response after surgery in two independent cohorts of patients with resectable, stage II–III gastric cancer. The single patient classifiers could complement TNM staging to optimise decision making in patients with resectable gastric cancer who are eligible for adjuvant chemotherapy after surgery. Further validation of these results in prospective studies is warranted.

Funding Ministry of ICT and Future Planning; Ministry of Trade, Industry, and Energy; and Ministry of Health and Welfare.

Introduction

The current standard of care for resectable, stage II–III gastric cancer includes adjuvant chemotherapy after surgery to prevent disease recurrence and improve survival.[1–4] The use of this regimen is largely based on the results of the CLASSIC trial (NCT00411229),[4] which reported a survival improvement in patients with gastric cancer treated with adjuvant capecitabine plus oxaliplatin after primary gastrectomy with D2 lymph node dissection compared with patients who received surgery alone. However, in the 5-year follow-up analysis of the CLASSSIC trial, overall survival in the

Lancet Oncol 2018

Published Online
March 19, 2018
http://dx.doi.org/10.1016/S1470-2045(18)30108-6

See Online/Comment
http://dx.doi.org/10.1016/S1470-2045(18)30131-1

*Contributed equally

Department of Surgery, Yonsei University Health System (Prof J-H Cheong MD, Y Y Choi MD, Prof W J Hyung MD, Prof S H Noh MD), **Department of Biochemistry & Molecular Biology** (Prof J-H Cheong, Prof Y-M Huh MD), **Department of Pathology** (Hyu Kim MD), **Biostatistics Collaboration Unit** (H Y Kim MS, J Lee PhD), **and Department of Radiology** (Prof Y-M Huh), **Yonsei University College of Medicine, Seoul, South Korea; Yonsei Biomedical Research Institute, Seoul, South Korea** (Prof J-H Cheong, Prof Y-M Huh); **Brain Korea 21 PLUS Project for Medical Science, Seoul, South Korea** (Prof J-H Cheong); **Department of Surgery and Cancer Research Institute** (Prof H-K Yang MD) **and Department of Pathology** (Prof W H Kim MD), **Seoul National University College of Medicine, Seoul, South Korea; Department of Cancer Control and Population Health, Graduate School of Cancer Science and Policy** (Y-W Kim MD) **and Center for Gastric Cancer** (Y-W Kim, M-C Kook MD), **National Cancer Center, Goyang, South Korea; Department of Surgery, Chonnam National University Hwasun Hospital, Gwangju, South Korea** (Prof Y-K Park MD); **Department of Surgery** (Prof H-H Kim MD) **and Department of Pathology** (Prof H S Lee MD), **Seoul National University Bundang Hospital, Seongnam, South Korea; Department of Internal Medicine** (Prof K H Lee MD) **and**

Department of Pathology (Prof M J Gu MD), **Yeungnam University College of Medicine, Daegu, South Korea; Department of Surgery** (Prof S H Choi MD) **and Department of Pathology** (Prof S Hong MD), **Gangnam Severance Hospital, Seoul, South Korea; Department of Surgery, Chung-Ang University Hospital, Chung-Ang University College of Medicine, Seoul, South Korea** (J W Kim MD); **MediBio-Informatics Research Center, Novomics, Seoul, South Korea** (E Jang PhD, Hye Kim MS, Prof Y-M Huh); **and YUHS-KRIBB Medical Convergence Research Institute, Seoul, South Korea** (Prof Y-M Huh)

Correspondence to:
Prof Sung Hoon Noh, Department of Surgery, Yonsei University College of Medicine, Seoul 03722, South Korea
sunghoonn@yuhs.ac

Research in context

Evidence before this study

Gastric cancer is one of the most frequently diagnosed cancers and the third leading cause of cancer-related mortality globally. Adjuvant chemotherapy after D2 gastrectomy was shown to be beneficial compared with D2 gastrectomy alone for patients with resectable gastric cancer in two phase 3, randomised controlled trials (the ACTS-GC and CLASSIC trials), and was recommended as the standard of care in this setting by the National Comprehensive Cancer Network guidelines in 2012. We searched PubMed before the study start for articles published between Jan 1, 2001, and Dec 30, 2011, without language restrictions, using the terms ["gastric cancer" AND "adjuvant chemotherapy" AND "benefit"] AND ["predictive" OR "marker" OR "subtype" OR "sensitive"]. We excluded reviews, studies of prognostic markers, and those with irrelevant eligibility criteria. We identified one retrospective study with a small, single cohort. We updated our search for articles published between Jan 1, 2012, and Aug 31, 2017, using the same criteria. Among the eight studies found, two developed their predictive model on the basis of marker discovery strategies. The remaining six studies selected biomarkers on the basis of knowledge gained from previous studies and evaluated their functions or clinical implications. No study investigated independent cohorts of patients with gastric cancer to assess the clinical validity of their predictive models or markers, or reported a test with predictive chemotherapy response for patients with resected gastric cancer that has been validated.

Added value of this study

We show that a predictive single patient classifier test based on gene expression analysis can identify patients with resectable gastric cancer who will benefit from adjuvant chemotherapy after D2 gastrectomy, independently of staging and risk-stratification information. To the best of our knowledge, this study is the first to show a significant association between a predictive test and patients' response to adjuvant chemotherapy in two independent cohorts of patients with gastric cancer.

Implications of all the available evidence

Stratification of patients into immune, stem-like, and epithelial subtypes with the single patient classifers described here will help to identify the subset of patients with localised gastric cancer who might benefit from adjuvant therapy after surgery. The predictive single patient classifier has the potential to inform molecularly motivated and patient subtype-oriented therapeutic decisions for patients with stage II–III, resectable gastric cancer.

surgery-only group was 69% (95% CI 64–73) compared with 78% (74–82) in the adjuvant chemotherapy group, with a moderate absolute benefit from adjuvant chemotherapy of 9%. These findings highlight the need for a predictive test that identifies patients with resectable gastric cancer at different risk of recurrence and determines the likelihood of chemotherapy benefit.

With the advent of high-throughput technologies, large sets of tumour genomic data have been obtained and systematically analysed to understand the biological complexity of cancer phenotypes. Biomarkers have been developed and used to identify patients with particular cancer phenotypes of interest; for example, a genomic test with clinical utility in predicting chemotherapy benefit for patients with breast cancer has been validated and commercialised.[5,6] In colon cancer, some molecular subtypes have been shown to have predictive value for chemotherapy benefit and prognostic relevance.[7–10] Although several studies[11–18] have identified molecular subtypes of gastric cancer associated with prognosis, because of the absence of independent and relevant patient cohorts, molecular subtypes useful for individual risk assessment and for predicting adjuvant chemotherapy benefit have not been yet validated.

Because of the partial overlap between gastric and colon cancer in terms of molecular subtypes and clinical characteristics (eg, in both tumours, microsatellite instability was associated with improved clinical outcomes when chemotherapy was withheld,[11,14,19,20] a specific molecular subtype [enterocyte[7,8] or canonical[10]] was associated with a proliferative signature and chemotherapy response,[12,18] and an epithelial–mesenchymal transition [EMT] signature was associated with a poor prognosis[8,13]), we postulated that at least three molecularly definable, clinical subsets of stage II–III gastric cancer exist, including low-risk patients who might have a good prognosis when treated with surgery alone, and patients with poor prognosis, among whom some might benefit from adjuvant chemotherapy and some might not. The modest benefit observed in patients with stage II–III gastric cancer treated with adjuvant chemotherapy after surgery in the CLASSIC trial[4] suggests that the absolute benefit of adjuvant chemotherapy in this setting could be maximised if patient subsets are defined in terms of chemotherapy response likelihood. Therefore, in this retrospective study we aimed to develop and validate a single patient classifier assay that could assign patients to subsets with predictive value for chemotherapy response on the basis of gene-expression data obtained from resected, formalin-fixed, paraffin-embedded (FFPE) tumour tissues.

Methods

Study design and data sources

In this multi-cohort, retrospective study, we used a multi-step strategy to develop and validate gene expression-based, single patient classifiers to stratify patients with gastric cancer according to risk parameters

(prognostic single patient classifier) and identify patients who would benefit from adjuvant chemotherapy after D2 gastrectomy (predictive single patient classifier). The single patient classifiers based on a real-time RT-PCR assay were designed using several genomic datasets from patients with gastric cancer (exploratory study), an internal cohort of patients who received D2 gastrectomy plus adjuvant chemotherapy or D2 gastrectomy alone (discovery and evaluation study), and an external validation cohort treated with D2 gastrectomy plus adjuvant chemotherapy or D2 gastrectomy alone as part of the CLASSIC trial (figure 1).

To identify potential candidates for the single patient classifier genes during the exploratory phase of the study, we used five cohorts of patients with gastric cancer with transcriptome profiling data available (n=1259 tumour samples); four from GEO DataSets, including one gene expression dataset from a cohort available at the Yonsei Cancer Center (GSE15459 [n=200],[11] GSE62254 [n=300],[13] GSE13861 [n=64],[14] and GSE84437 [n=433] internal dataset), and one from the TCGA Data Portal (n=262[12]). We selected these exploratory cohorts because of the large sample size, quality of the transcriptome data evaluable (RNA extracted from fresh, non-fixed, tumour samples), and availability of the gene expression-based molecular classification and clinical characteristics of patients, including some information about treatments (appendix pp 1, 20). Treatment information for the exploratory cohorts was largely unavailable because these were treated before 2010, when, because of the absence of standard guidelines for adjuvant chemotherapy in patients with stage II–III gastric cancer, different regimens or doses were administered as per the treating physician's discretion.

For the discovery and evaluation of the single patient classifier assay, we used FFPE tumour tissues from a cohort of 310 patients with resectable, stage II–III gastric cancer who were treated at the Yonsei Cancer Center (Seoul, South Korea) between Jan 1, 2005, and Dec 31, 2010, including 193 specimens from patients treated with D2 gastrectomy followed by fluorouracil-based adjuvant chemotherapy and 117 specimens from patients treated with surgery alone.

For the external validation of the single patient classifiers, we used a selection of archived FFPE tumour tissues from patients with resectable, stage II–III gastric cancer who were part of the multicentre, randomised, open-label, phase 3 CLASSIC trial[4] undertaken in 37 centres in South Korea, China, and Taiwan. Of the 1035 patients with gastric cancer randomised in the CLASSIC trial to receive D2 gastrectomy plus adjuvant chemotherapy (eight 3-week cycles of oral capecitabine [1000 mg/m²] twice a day on days 1–14 plus intravenous oxaliplatin [130 mg/m²] on day 1) or surgery only between June, 2006, and June, 2009, 629 tumour specimens were obtained along with follow-up data from the high-accruing institutions to use in this retrospective analysis as an independent validation cohort.

Specimens from the validation cohort used for RNA extraction in this retrospective study were resected from patients aged 18 years or older who had stage II–IIIB gastric cancer with no evidence of metastatic disease. Patients were assessed for disease by abdominal CT, MRI, or chest radiograph scans. In this retrospective study, patients were staged according to the American Joint Committee on Cancer (AJCC) cancer staging manual (sixth edition).[21] The institutional review board (IRB) at each institution approved use of the procured tissues and waived the need for informed consent for this retrospective analysis of anonymous data.

Procedures

The multi-step strategy included identifying candidate genes in gastric cancer transcriptome datasets (exploratory study), selecting genes compatible with clinical grade

For more on **GEO DataSets** see https://www.ncbi.nlm.nih.gov/gds

See **Online** for appendix

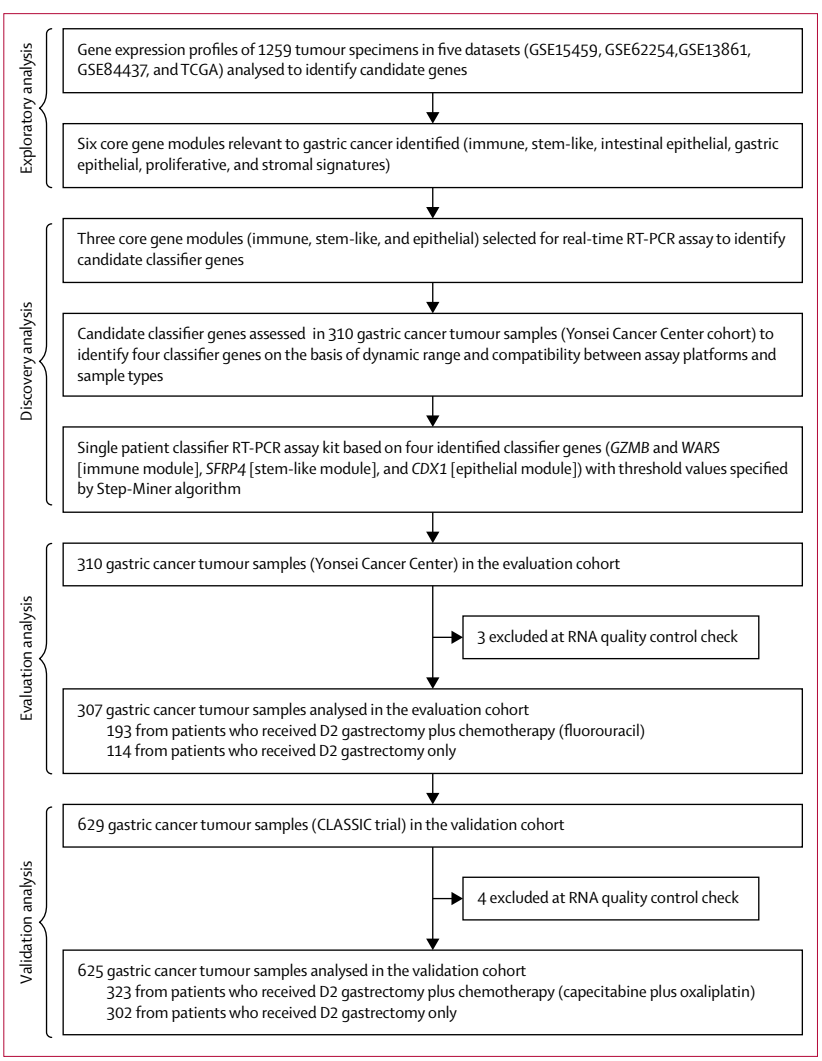

Figure 1: **Outline of the overall study flow**

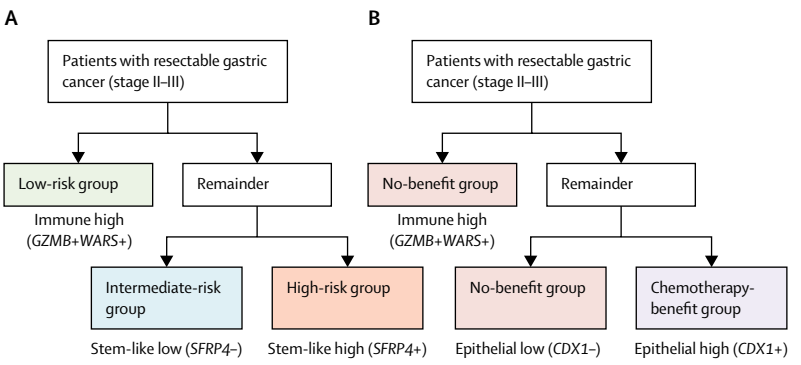

Figure 2: **Clinical subsets of patients with resectable gastric cancer by prognostic (A) or predictive (B) single patient classifiers**
Patients are assigned to the different subgroups on the basis of expression of four classifier genes assessed with real-time RT-PCR.

	Yonsei Cancer Center (discovery and evaluation cohort; n=310)	Collected tumour samples from patients in the CLASSIC trial (validation cohort; n=629)
Age (years)		
<65	187 (60·3%)	450 (71·5%)
≥65	123 (39·7%)	179 (28·5%)
Sex		
Female	91 (29·4%)	190 (30·2%)
Male	219 (70·6%)	439 (69·8%)
TNM stage*		
II	143 (46·1%)	303 (48·2%)
III	167 (53·9%)	326 (51·8%)
T status		
T1T2	135 (43·5%)	352 (56·0%)
T3T4	175 (56·5%)	277 (44·0%)
N status		
N0	38 (12·3%)	52 (8·3%)
N1N2	272 (87·7%)	577 (91·7%)
Adjuvant chemotherapy		
Yes	193 (62·3%)	325 (51·7%)
No	117 (37·7%)	304 (48·3%)

Data are n (%). TNM=tumour–node–metastasis. T status=primary tumour status. N status=nodal status. *Defined according to the American Joint Committee on Cancer sixth edition.

Table 1: **Clinical characteristics of patients in the evaluation and validation cohorts**

assays to define the single patient classifiers relevant to stratify patients with resectable gastric cancer (discovery study), assessing the clinical validity of the single patient classifiers in specimens from an internal cohort (evaluation study), and external validation of the results using an independent cohort (validation study; figure 1 and appendix p 8).

In the exploratory study, we used bioinformatics algorithms (unsupervised clustering analyses, followed by gene-set enrichment analyses)[7,9,22,23] to identify candidate genes for the single patient classifiers based on their association with patients' outcomes (prognostic single patient classifier) or chemotherapy response (predictive single patient classifier; see appendix p 2 for more details). The exploratory study did not adhere to a prespecified protocol, but followed the guidelines for biomarker development suggested by the US Institute of Medicine.[24] In the discovery study, we developed a real-time RT-PCR assay to screen the identified candidate genes in search of the genes with relevant biological function in gastric cancer and robust analytical performance parameters for quantitative PCR testing defined in the appendix (p 4). Total RNA was extracted from FFPE tumour samples from the Yonsei Cancer Center institutional tumour bank using MasterPure Complete DNA and RNA Purification Kit (Epicentre Technologies, Madison, WI, USA), and real-time RT-PCR was done with the SensiFAST Probe Lo-ROX Kit (Bioline, London, UK). The expression threshold required for candidate genes to be considered for the single patient classifiers was established with a commercial Good Manufacturing Practice (GMP)-grade kit (nProfiler I; Novomics, Seoul, South Korea) using the samples included in the discovery cohort (appendix p 5). For quality control, the A260:A280 ratio of the extracted RNA had to be above 1·8 and the A260:A230 ratio had to be above 1·6. The upper and lower limits of quantitation for quantitation cycle (Cq) values obtained with the nProfiler I assay are provided in the appendix (p 21).

In the evaluation phase, patients from the Yonsei Cancer Center evaluation cohort were stratified according to the proposed single patient classifiers selected in the discovery phase, and the associations with chemotherapy response by treatment-interaction design (predictive value) and overall survival by Cox proportional hazard model (prognostic value) were analysed. Once the optimal threshold levels for the individual genes included in each single patient classifier were determined, and their potential clinical utility confirmed by interaction tests, the single patient classifiers were defined.

In the validation study, the prognostic and predictive chemotherapy response value of the defined single patient classifiers was tested in an independent validation cohort of patients with gastric cancer. RNA was extracted from FFPE tumour tissues from patients who were treated as part of the CLASSIC trial. According to the IRB-approved protocol of the CLASSIC trial (NCT03403296), and following Simon and colleagues' principle of "prospective–retrospective designs using archived specimens for evaluation of prognostic and predictive biomarkers",[25] individuals who did the real-time RT-PCR and laboratory analyses were masked to the clinical data; only the two independent statisticians (HYK and JL) who did the validation analyses had access to the clinical data (data handling process detailed in appendix p 10).

Statistical analysis

In the exploratory study, we used a nearest shrunken centroid method and gene-set enrichment analysis to characterise clustered features from the gastric cancer transcriptome.[7,8,22,23] A Step-Miner regression model was used to establish cutoff values for each classifier gene identified in the discovery study.[26] A χ^2 test was used to assess whether clinical variables of the collected CLASSIC trial specimens were balanced compared with the entire CLASSIC cohort and to test whether or not the variables were associated with the prognostic and predictive single patient classifiers. We assessed 5-year overall survival, defined as the time from randomisation to death from any cause, in the evaluation (Yonsei Cancer Center) and validation (CLASSIC) cohorts using the Kaplan-Meier method, log-rank tests, the Cox proportional-hazards model with Wald statistics, and interaction tests. The multivariate Cox analysis was adjusted for age, sex, T and N statuses, and chemotherapy. We calculated 95% CIs using standard normal distribution.

Because the validation analysis had a retrospective design, the number of samples available for analysis was fixed (we included all patients in the CLASSIC trial with available FFPE tissues). We estimated that, with the samples available, we would achieve 80% power in the interaction analysis of the predictive single patient classifier and chemotherapy response outcomes if the range of ratios of hazard ratios (HRs) between the chemotherapy-benefit and no-benefit groups was set as 0·4–0·6 (appendix pp 7, 17). This estimation was done with the SWOG web tool.

Two-sided p values are reported, and the significance level was set to less than 0·05. All statistical analyses, except for the power analysis, were done with RGUI, version 3.3.3.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. J-HC and SHN had full access to all

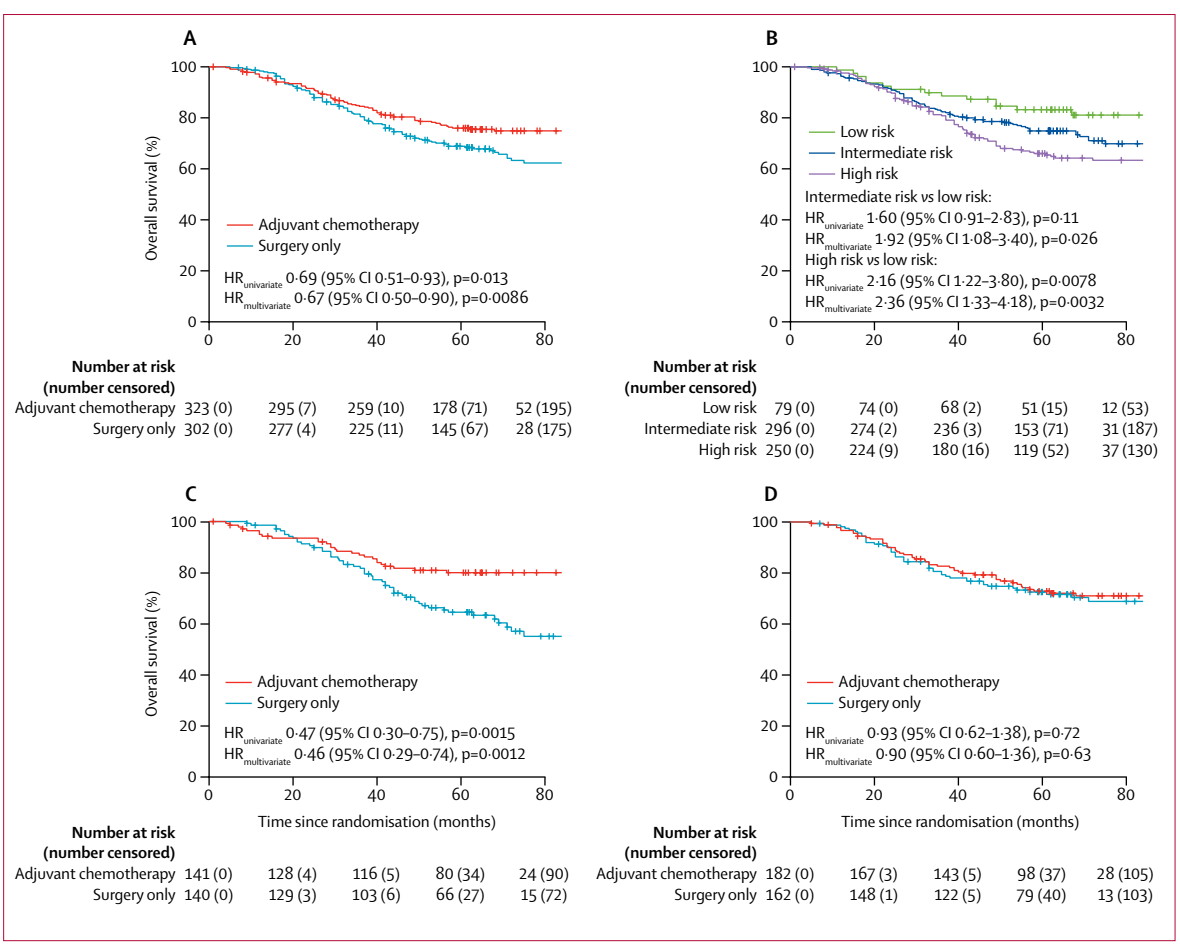

Figure 3: **Overall survival in the validation cohort by single patient classifiers**

625 of the 629 tumour samples from patients in the CLASSIC trial are included in these analyses; four samples were excluded during the RNA quality control evaluation. (A) Overall survival by treatment (D2 gastrectomy plus adjuvant chemotherapy or D2 gastrectomy only). (B) Overall survival by prognostic single patient classifier groups. (C) Overall survival by predictive single patient classifier, chemotherapy-benefit group, and treatment received. (D) Overall survival by predictive single patient classifier, no-benefit group, and treatment received. HR=hazard ratio.

data in the study and as corresponding authors had final responsibility for the decision to submit for publication.

Results

In the exploratory study, we analysed the gene-expression profiles of 1259 gastric cancer tumour specimens with patient clinical data available to identify key biological processes underlying potential molecular subtypes predictive of chemotherapy response and prognosis in patients with gastric cancer (figure 1; appendix p 8). First, we used bioinformatics algorithms to identify five molecular subtypes (inflammatory, intestinal, gastric, mixed-stromal, and mesenchymal) with different prognoses in terms of overall survival (appendix p 11) and 34 gene modules, of which 21 were closely associated with the subtypes and clustered into six biological categories with relevance in gastric cancer (immune, stem-like, intestinal epithelial, gastric epithelial, proliferative, and stromal signatures; appendix p 11). From these six core gene modules, we selected the immune and stem-like signatures as prognostic classifiers because of their association with good and poor clinical outcomes, respectively, in previous studies.[13,18] For the predictive chemotherapy-benefit classifiers, we hypothesised that the proliferation gene module could predict chemotherapy response based on similar results obtained in patients with breast and colon cancer.[5,6,8] However, the narrow difference in the expression level of proliferation genes and the fact that these genes are less subtype specific, could have increased classification errors. Thus, instead, we selected the intestinal epithelial gene module as a surrogate classifier for the proliferation gene module, on the basis of its systematic correlation with cell cycle-associated genes and because it has a wider dynamic range to act as a classifier. Collectively, we selected three gene modules whose member genes could be potential classifiers for a predictive test: immune, stem-like, and epithelial.

To translate the identified three gene modules into a clinical-grade test, in the discovery study, we designed real-time RT-PCR probes for 239 genes representing the three gene modules (48 genes for the immune module, 131 for the stem-like module, and 60 for the intestinal epithelial module) and assessed their compatibility between platforms and sample types; this analysis returned 12 candidate genes. We analysed the expression of these 12 genes regarding their dynamic range and intra-tumour heterogeneity in 310 institutional FFPE tumour tissues (Yonsei Cancer Center cohort) with real-time RT-PCR, and selected four classifier genes (*GZMB* and *WARS* representing the immune module, *SFRP4* representing the stem-like module, and *CDX1* representing the intestinal epithelial module). Five genes were identified as internal standards to normalise the quantitative PCR analyses (*ACTB*, *ATP5E*, *HPRT1*, *GPX1*, and *UBB*). Using the real-time RT-PCR probes for the nine genes, we establish a GMP-grade single patient classifier assay kit.

We used Step-Miner[26] to establish the cutoff point (threshold ΔCq values) for each classifier gene to stratify patients as positive or negative according to the predictive (*CDX1*, intestinal epithelial module) and prognostic (*GZMB* and *WARS*, immune module; and *SFRP4* stem-like module) single patient classifiers. This analysis was done using 307 FFPE gastric cancer samples from the Yonsei Cancer Center cohort (figure 1); three of 310 tissue samples were excluded from the assay because they failed the RNA quality control test. The established cutoff points for each gene were −5·18 for *GZMB*, −2·14 for *WARS*, −3·63 for *SFRP4*, and −2·69 for *CDX1* (appendix p 5). Using these cutoff points, two rule-based single patient classifier algorithms were developed, and patients were classified into low-risk (immune-high), intermediate-risk (immune-low and stem-like-low), or high-risk (immune-low and stem-like-high) groups according to the prognostic single patient classifier (figure 2A) and into no-benefit (immune-high or immune-low and epithelial-low) or chemotherapy-benefit (immune-low and epithelial-high) groups according to the predictive single patient classifier (figure 2B).

We evaluated the clinical validity of the single patient classifiers in the specimens from 307 patients in the Yonsei Cancer Center cohort, including 193 treated with D2 surgery followed by fluorouracil-based chemotherapy and 114 treated with surgery alone (figure 1). Median follow-up for the 307 patients was 59 months (IQR 36–72) for overall survival. The prognostic single patient classifier

	n (%)	5-year overall survival*	HR (95% CI); p value	
			Univariate	Multivariate†
Low risk	79 (13%)	83·2% (75·2–92·0)	1 (ref)	1 (ref)
Intermediate risk	296 (47%)	74·8% (69·9–80·1)	1·60 (0·91–2·83); 0·11	1·92 (1·08–3·41); 0·026
High risk	250 (40%)	66·0% (60·1–72·4)	2·16 (1·22–3·80); 0·0078	2·36 (1·33–4·18); 0·0032

HR=hazard ratio. *Data are % (95% CI). †Model adjusted for age, sex, primary tumour and nodal statuses, and previous chemotherapy.

Table 2: **5-year overall survival in the validation cohort by prognostic single patient classifier groups**

	n (%)	5-year overall survival*		HR (95% CI); p value†	
		Surgery plus adjuvant chemotherapy	Surgery only	Univariate	Multivariate‡
Chemotherapy benefit	281 (45%)	80·0% (73·5–87·1)	64·5% (56·8–73·3)	0·47 (0·30–0·75); 0·0015	0·46 (0·29–0·74); 0·0012
No benefit	344 (55%)	72·9% (66·5–79·9)	72·5% (65·8–79·9)	0·93 (0·62–1·38); 0·71	0·90 (0·60–1·36); 0·63

HR=hazard ratio. *Data are % (95% CI). †Calculated with the surgery-only group as the reference. ‡Model adjusted for age, sex, primary tumour and nodal statuses, and previous chemotherapy.

Table 3: **5-year overall survival in the validation cohort by predictive single patient classifier groups**

stratified 24 (8%) of 307 patients as low risk, 115 (37%) as intermediate risk, and 168 (55%) as high risk. 5-year overall survival was 83·3% (95% CI 69·7–99·7) in the low-risk group, 71·8% (63·9–80·6) in the intermediate-risk group, and 58·2% (51·1–66·3) in the high-risk group (p=0·0095; appendix pp 15, 23, 24). The predictive single patient classifier stratified 145 (47%) of 307 patients as potentially responsive to chemotherapy (chemotherapy-benefit group) and predicted 162 (53%) as having no benefit from chemotherapy (appendix p 25). The overall survival gain of adjuvant chemotherapy in the chemotherapy-benefit group versus the no-benefit group was significant in the multivariate analysis (HR 0·43 [95% CI 0·23–0·79] *vs* 1·20 [0·66–2·19], p=0·0066), but not in the univariate analysis (0·70 [0·41–1·22] *vs* 1·09 [0·64–1·87], p=0·21; appendix pp 16, 26). The interaction between the predictive single patient classifier groups and adjuvant chemotherapy benefit was also significant after adjustment for age, sex, T status, and N status ($p_{interaction}$=0·039; appendix p 26). To validate the single patient classifiers, we collected 629 (61% of 1035 patients) archived FFPE tumour samples from the CLASSIC trial (figure 1, table 1). The study subset was representative of the entire CLASSIC trial cohort, including 1035 patients, in terms of clinical and pathological features, as well as chemotherapy benefit (appendix p 27). Four of the 629 collected specimens were excluded per quality-control criteria, and 625 samples were evaluated using the single patient classifier assay, including 323 specimens from patients treated with adjuvant chemotherapy post-surgery and 302 from patients who received surgery only (figure 3A). Median follow-up was 61 months (IQR 43–74) for overall survival. The prognostic single patient classifier assigned 79 (13%) of 625 tumour samples to the low-risk group, 296 (47%) to the intermediate-risk group, and 250 (40%) to the high-risk group (table 2; appendix p 30). 5-year overall survival was 83·2% (95% CI 75·2–92·0) in the low-risk group, 74·8% (69·9–80·1) in the intermediate-risk group, and 66·0% (60·1–72·4) in the high-risk group (p=0·012; figure 3B, table 2). 5-year overall survival in the low-risk group was significantly improved compared with that of the high-risk group in both univariate and multivariate regression analyses (table 2; appendix p 18). The predictive single patient classifier assigned 281 (45%) of 625 patients to the chemotherapy-benefit group and 344 (55%) to the no-benefit group (table 3; appendix p 31). Of the 344 patients assigned to the no-benefit group, 79 (23%) were immune-high and 265 (77%) were immune-low and epithelial-low. Patients classified by the single patient classifier in the chemotherapy-benefit group benefited from adjuvant chemotherapy compared with those who received surgery alone; 5-year overall survival was 64·5% (95% CI 56·8–73·3) in patients who received surgery only versus 80·0% (73·5–87·1) in those who received adjuvant chemotherapy (univariate HR 0·47 [95% CI 0·30–0·75], p=0·0015; multivariate HR 0·46 [95% CI 0·29–0·74], p=0·0012; figure 3C, table 3). By contrast, among patients assigned to the no-benefit group by the single patient classifier, 5-year overall survival was similar between patients who received adjuvant chemotherapy and those who did not (figure 3D, table 3).

The association between the predictive single patient classifier groups (chemotherapy-benefit *vs* no-benefit) and chemotherapy response in terms of overall survival was significant ($p_{interaction}$=0·036 in the univariate analysis and $p_{interaction}$=0·048 in the multivariate analysis; figure 4). There was no association between the predictive single patient classifier groups and T status (data not shown;

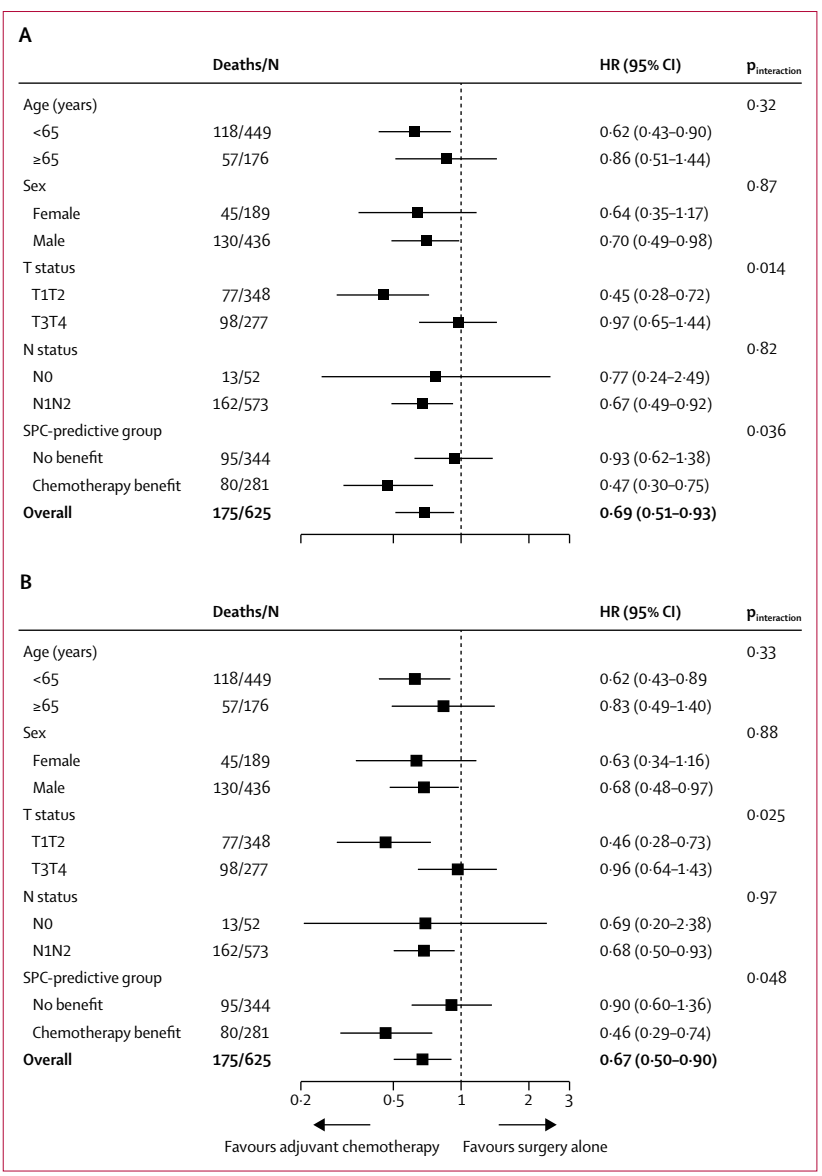

Figure 4: **Association between overall survival and predictive single patient classifier groups or baseline characteristics in the validation cohort**

(A) Univariate analysis. (B) Multivariate analysis adjusted for age, sex, T and N statuses, and SPC-predictive groups. p values for association between clinical variables and adjuvant chemotherapy benefit are shown. T status=primary tumour status. N status=nodal status. SPC=single patient classifer.

$p_{interaction}$=0·60 in the univariate analysis and $p_{interaction}$=0·65 in the multivariate analysis).

Discussion

In this study, we developed a clinical-grade, four-gene test with prognostic value and predictive chemotherapy response benefit for patients with resectable, stage II–III gastric cancer. The validated classifier genes were *GZMB*, *WARS*, *SFRP4*, and *CDX1*, which were selected for their analytical robustness in a real-time RT-PCR assay and biological relevance in gastric cancer. *GZMB* (granzyme B) and *WARS* (tryptophanyl-tRNA synthetase) are related to immune regulation and inflammatory response; *SFRP4* (secreted frizzled-related protein 4) is a *WNT* signalling-associated EMT modulator; and *CDX1* (caudal-type homeobox 1) is a biomarker for gastric intestinal metaplasia, an intermediate, precancerous lesion in gastric carcinogenesis. We used a collection of tumour samples from the CLASSIC trial to confirm that the single patient classifiers can predict prognosis and chemotherapy benefit in heterogeneous cohorts of patients with resectable gastric cancer. The predictive single patient classifier assigned a subset of the CLASSIC patients to the chemotherapy-benefit group. These patients were shown to have improved overall survival after chemotherapy compared with patients who only received surgery, with a mean absolute increase in 5-year overall survival of 15·5%. By contrast, patients assigned by the predictive single patient classifier to the no-benefit group did not benefit from chemotherapy, despite having a poor prognosis. Although this analysis was retrospective, we used archived specimens from a randomised trial to validate the results and followed a prospectively designed and retrospectively tested approach;[25] the validation cohort was well balanced to represent the original cohort in the CLASSIC trial, the assay was analytically validated and done blinded, and the validation analysis was prespecified (NCT03403296; appendix pp 6, 16).

The treatment effect by predictive single patient classifier groups in the evaluation (Yonsei Cancer Center) cohort was not significantly associated with chemotherapy response benefit in the univariate analysis. This finding might be due to the absence of a randomised, controlled clinical setting during treatment; in the chemotherapy-benefit group, adjuvant chemotherapy was frequently withheld from patients with early T-status tumours (appendix pp 6, 25). By contrast, the prospective validation cohort (CLASSIC trial) enabled assessment of the clinical validity of the predictive test, and the contradictory results in the univariate analysis indicate that further validation of these results in prospectively treated cohorts is warranted.

Although previous studies[11–18] have also used gene-expression profiling data to develop predictors for chemotherapy response benefit and prognosis in patients with gastric cancer, none have validated the results because of several intrinsic problems associated with the study designs used (small sample sizes that could not reflect the heterogeneity of gastric cancer, heterogeneous treatment regimens received by patients, different assay platforms, and suboptimal statistical analysis approaches). Validation of the results of this study with the CLASSIC trial cohort was achieved through use of a large-scale, data-driven approach, allowing the development of single patient classifiers with biological and clinical relevance for patients with gastric cancer. Additionally, the classification of immune-high patients in both the prognostic and predictive single patient classifiers might have reduced the overall heterogeneity of the cohort because immune-high tumours can indicate gastric cancer with Epstein-Barr virus infection or deficient DNA mismatch repair, which are aetiologically distinct from other gastric cancer subtypes.[27] Moreover, the single patient classifier assay was based on real-time RT-PCR, which is generally considered the gold-standard method for RNA quantitation. Although a concern might be that use of FFPE tumour tissues decreases the reliability of RNA quantitation, technical efforts in the real-time RT-PCR procedure, including the design of optimal primers and amplicons, addressed this challenge. A real-time RT-PCR-based predictive test using FFPE breast cancer tissues has been used in routine clinical practice in the USA.[28]

Clinically, the single patient classifier assay provides clinicians with auxiliary information distinct from the standard risk-stratification methods, such as TNM staging and histological grade, which might be useful when making decisions about adjuvant therapy after D2 surgery for patients with stage II–III gastric cancer. The prognostic single patient classifier not only stratifies patients into groups with different long-term outcomes, but also informs on the biological characteristics of the tumours, complementing the anatomy-based staging system. Patients with stage II gastric cancer, whose prognoses were generally expected to be favourable, were slightly more enriched in the intermediate-risk group than in the other risk groups (appendix p 30), suggesting that the prognostic single patient classifier provides clinically important information independent of TNM staging. Furthermore, dual assignment of individual patients to the prognostic and predictive single patient classifier groups offers better postoperative management based on actionable information than the currently used TNM staging system. In particular, adjuvant chemotherapy could be withheld from patients assigned to the immune-high subgroup in the no-benefit group (these patients have the most favourable prognosis with surgery alone) to avoid unnecessary adverse effects related to chemotherapy.[9,13,19,20] Additionally, for patients assigned to the chemotherapy-benefit group, adjuvant chemotherapy is strongly recommended because their intermediate-risk prognosis is the result of a sizeable chemotherapy benefit. Lastly, for patients in the

no-benefit group classed as immune-low and epithelial-low, whose response to adjuvant chemotherapy is uncertain, there is an unmet clinical need to find new therapeutic options. The absence of chemotherapy benefit in patients classified as immune-low and epithelial-low might reflect the biological nature of their tumour cells (mesenchymal tumour cells with low proliferation). In colon cancer, the mesenchymal or EMT subtype was associated with absence of benefit from chemotherapy,[8] confirming our findings in patients with immune-low and epithelial-low tumours.

Our study has some limitations. First, the exploratory and discovery analyses were done without a prespecified protocol. Second, although the clinical variables that are known to affect patients' outcomes were well balanced between the validation cohort (n=629) and the original CLASSIC trial cohort (n=1035), the collection of specimens for the validation cohort could have introduced selection bias. Second, archived specimens from the phase 3, randomised ACTS-GC trial (NCT00152217)[3] could not be used as an additional validation cohort because the specimen storage period had expired. Third, the single patient classifiers were validated in cohorts including mostly Asian patients with stage II–III gastric cancer, and therefore validation in non-Asian patients is warranted to validate the generalisability of the results. Additionally, validation is necessary to confirm the reproducibility of these results and the generalisability of the single patient classifiers in prospectively designed studies that incorporate other races, regions, cancer stages, or interventions (eg, S1-based chemotherapy regimens). Clinical studies that address potential strategies for clinical care based on single patient classifier information would also be relevant. For example, prospective trials could stratify patients with stage II–III, resectable gastric cancer by the single patient classifiers to test novel therapeutics, or apply the single patient classifiers to different settings using data from randomised trials (eg, the MAGIC trial,[2] which evaluated perioperative chemotherapy in non-Asian patients with resectable gastric cancer. Moreover, the applicability of the test in the preoperative or perioperative setting could be evaluated in prospectively designed, multi-country trials. We expect that the single patient classifiers described in this study will be of use in other clinically relevant settings, such as in decision making on neoadjuvant therapy in patients with gastric cancer with available endoscopic biopsy samples.

In conclusion, the single patient classifiers described in this study for the stratification of patients with localised gastric cancer into immune, stem-like, and epithelial subtypes has the potential to inform patient selection for adjuvant therapy after surgery. Our predictive single patient classifier could inform decision making beyond the postoperative setting. Further validation of these results in prospective studies is warranted.

Contributors

J-HC and SHN conceived, designed, and supervised the study. J-HC, H-KY, HyuK, WHK, Y-WK, M-CK, Y-KP, H-HK, HSL, KHL, MJG, SHC, SH, JWK, YYC, WJH, and SHN collected the specimens and constructed the database. HyuK offered pathological review. HYK and JL did the statistical analysis. J-HC, EJ, and Y-MH wrote and revised the manuscript. EJ and Y-MH did the computational analysis. HyeK and Y-MH did the real-time RT-PCR assay. All authors had full access to the study data, discussed and reviewed the manuscript, and approved the manuscript for publication.

Declaration of interests

To develop a commercial-grade real-time RT-PCR assay that was fully analytically validated before being applied to the archived tumour tissues from the CLASSIC trial, as required by the US Food & Drug Administration guideline for multigene prognostic assays, we founded NOVOMICS. J-HC, SHN, and Y-MH are founders and shareholders of NOVOMICS. J-HC, SHN, HyuK, and Y-MH applied for a patent. All other authors declare no competing interests.

Acknowledgments

This study was supported by the Bio & Medical Technology Development Program of the National Research Foundation funded by the Ministry of ICT and Future Planning (NRF-2015M3A9D7029834); the Technology Innovation Program funded by the Ministry of Trade, Industry, and Energy (10041653); the National R&D Program for Cancer Control funded by the Ministry of Health and Welfare (1020390 and 1320360); and the Korea Health Technology R&D Project through the Korea Health Industry Development Institute funded by the Ministry of Health and Welfare (HI14C1324). We thank Soonmyung Paik (Severance Biomedical Science Institute, Yonsei University College of Medicine, Seoul, South Korea) for his invaluable comments and advice.

References

1 National Comprehensive Cancer Network. Clinical practice guidelines in oncology. Gastric cancer. Version 5. 2017. https://www.nccn.org/professionals/physician_gls/pdf/gastric.pdf (accessed Dec 17, 2017).

2 Cunningham D, Allum WH, Stenning SP, et al. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *N Engl J Med* 2006; **355:** 11–20.

3 Sasako M, Sakuramoto S, Katai H, 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.

4 Noh SH, Park SR, Yang HK, 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.

5 Paik S, Tang G, Shak S, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor–positive breast cancer. *J Clin Oncol* 2006; **24:** 3726–34.

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

7 Sadanandam A, Lyssiotis CA, Homicsko K, et al. A colorectal cancer classification system that associates cellular phenotype and responses to therapy. *Nat Med* 2013; **19:** 619–25.

8 Song N, Pogue-Geile KL, Gavin PG, et al. Clinical outcome from oxaliplatin treatment in stage II/III colon cancer according to intrinsic subtypes secondary analysis of NSABP C-07/NRG oncology randomized clinical trial. *JAMA Oncol* 2016; **2:** 1162–69.

9 Sargent DJ, Marsoni S, Monges G, et al. Defective mismatch repair as a predictive marker for lack of efficacy of fluorouracil-based adjuvant therapy in colon cancer. *J Clin Oncol* 2010; **28:** 3219–26.

10 Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med* 2015; **21:** 1350–56.

11 Lei Z, Tan IB, Das K, et al. Identification of molecular subtypes of gastric cancer with different responses to PI3-kinase inhibitors and 5-fluorouracil. *Gastroenterology* 2013; **145:** 554–65.

12 Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* 2014; **513:** 202–09.

13 Cristescu R, Lee J, Nebozhyn M, et al. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nat Med* 2015; **21:** 449–56.

14 Cho JY, Lim JY, Cheong JH, et al. Gene expression signature-based prognostic risk score in gastric cancer. *Clin Cancer Res* 2011; **17:** 1850–57.

15 Grau JJ, Domingo-Domenech J, Morente V, et al. Low thymidylate synthase expression in the primary tumor predicts favorable clinical outcome in resected gastric cancer patients treated with adjuvant tegafur. *Oncology* 2004; **66:** 226–33.

16 Cao Y, Liu H, Zhang H, et al. CXC chemokine receptor 1 predicts postoperative prognosis and chemotherapeutic benefits for TNM II and III resectable gastric cancer patients. *Oncotarget* 2017; **8:** 20328–39.

17 Jiang Y, Zhang Q, Hu Y, et al. ImmunoScore signature: a prognostic and predictive tool in gastric cancer. *Ann Surg* 2016; published online Dec 20. DOI:10.1097/SLA.0000000000002116.

18 Sohn BH, Hwang JE, Jang HJ, et al. Clinical significance of four molecular subtypes of gastric cancer identified by The Cancer Genome Atlas project. *Clin Cancer Res* 2017; **23:** 4441–49.

19 Kim SY, Choi YY, An JY, 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:** 819–25.

20 An JY, Kim H, Cheong JH, Hyung WJ, Kim H, Noh SH. Microsatellite instability in sporadic gastric cancer: its prognostic role and guidance for 5-FU based chemotherapy after R0 resection. *Int J Cancer* 2012; **131:** 505–11.

21 AJCC. AJCC Cancer staging manual, 6th edn. 2002. https://cancerstaging.org/references-tools/deskreferences/documents/ajcc6thedcancerstagingmanualpart1.pdf (accessed Feb 24, 2016).

22 Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci USA* 2004; **101:** 4164–69.

23 Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008; **9:** 559–71.

24 National Academies of Sciences, Engineering, and Medicine. Biomarker tests for molecularly targeted therapies: key to unlocking precision medicine. Washington, DC: The National Academies Press, 2016.

25 Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. *J Natl Cancer Inst* 2009; **101:** 1446–52.

26 Sahoo D, Dill DL, Tibshirani R, Plevritis SK. Extracting binary signals from microarray time-course data. *Nucleic Acids Res* 2007; **35:** 3705–12.

27 Chiaravalli AM, Feltri M, Bertolini V, et al. Intratumour T cells, their activation status and survival in gastric carcinomas characterised for microsatellite instability and Epstein-Barr virus infection. *Virchows Arch* 2006; **448:** 344–53.

28 National Comprehensive Cancer Network. Clinical practice guidelines in oncology. Breast cancer. Version 3. 2017. https://www.nccn.org/professionals/physician_gls/pdf/breast.pdf (accessed Dec 17, 2017).