Machine-learning model derived gene signature predictive of paclitaxel survival benefit in gastric cancer: results from the randomised phase III SAMIT trial

Raghav Sundar,1,2,3,4 Nesaretnam Barr Kumarakulasinghe,1 Yiong Huak Chan,5 Kazuhiro Yoshida,6 Takaki Yoshikawa,7 Yohei Miyagi,8 Yasushi Rino,9 Munetaka Masuda,9 Jia Guan,10 Junichi Sakamoto,11 Shiro Tanaka,10 Angie Lay-Keng Tan,2 Michal Marek Hoppe,12 Anand D. Jeyasekharan,1,12 Cedric Chuan Young Ng,13 Mark De Simone,14 Heike I. Grabsch 15,16 Jeeyun Lee,17 Takashi Oshima,18 Akira Tsuburaya,19 Patrick Tan20,21,22

INTRODUCTION

The identification of predictive biomarkers to personalise treatment of patients with gastric cancer (GC) is challenging due to the complex genomic and molecular landscape of this malignancy.1 Adding chemotherapy to surgery has improved patient with GC survival, when administered in either the perioperative or postoperative setting.2 Besides 5-fluorouracil (5FU) and

ABSTRACT

Objective To date, there are no predictive biomarkers to guide selection of patients with gastric cancer (GC) who benefit from paclitaxel. Stomach cancer Adjuvant Multi-Institutional group Trial (SAMIT) was a 2×2 factorial randomised phase III study in which patients with GC were randomised to Pac-S-1 (paclitaxel + S-1), Pac-UFT (paclitaxel + UFT), S-1 alone or UFT alone after curative surgery.

Design The primary objective of this study was to identify a gene signature that predicts survival benefit from paclitaxel chemotherapy in GC patients. SAMIT gene NanoString panel. A random forest machine-learning model was applied on the NanoString profiles to develop a gene signature. An independent cohort of metastatic patients with GC treated with paclitaxel and ramucirumab (Pac-Ram) served as an external validation cohort.

Results From the SAMIT trial 499 samples were analysed in this study. From the Pac-S-1 training cohort, the random forest model generated a 19-gene signature assigning patients to two groups: Pac-Sensitive and Pac-Resistant. In the Pac-UFT validation cohort, Pac-Sensitive patients exhibited a significant improvement in disease free survival (DFS): 3-year DFS 66% vs 40% (HR 0.44, p=0.0029). There was no survival difference between Pac-Sensitive and Pac-Resistant in the UFT or S-1 alone arms, test of interaction p>0.001. In the external Pac-Ram validation cohort, the signature predicted benefit for Pac-Sensitive (median PFS 147 days vs 112 days, HR 0.48, p=0.022).

Conclusion Using machine-learning techniques on one of the largest GC trials (SAMIT), we identify a gene signature representing the first predictive biomarker for paclitaxel benefit.

Trial registration number UMIN Clinical Trials Registry: C000000082 (SAMIT); ClinicalTrials.gov identifier: 02628951 (South Korean trial)

Significance of this study

What is already known on this subject?

► Paclitaxel is an active chemotherapeutic agent in the treatment of gastric cancer (GC).
► Stomach cancer Adjuvant Multi-Institutional group Trial (SAMIT) was a randomised 2×2 factorial phase III trial, studying the role of adjuvant paclitaxel in patients with GC.
► There are currently no biomarkers for selecting patients with GC who benefit from paclitaxel treatment.
► Machine-learning and genomic data is an emerging and novel strategy to develop reliable and valid predictive models.

What are the new findings?

► NanoString transcriptomic profiling was performed on SAMIT surgical resection samples.
► Using the trial’s 2×2 factorial study design, a gene signature predicting paclitaxel survival benefit was developed with a random forest machine-learning method.
► The model was validated internally within SAMIT samples, and independently in an external validation cohort.

How might it impact on clinical practice in the foreseeable future?

► These findings represent the first predictive biomarker for paclitaxel benefit in GC.
► Given the rising utilisation of taxanes in GC in the adjuvant and perioperative setting, this biomarker may guide in identifying patients who benefit from taxane based therapy, following further validation in prospective trials.

Stomach

Original research

Watch Video

Check for updates

© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.
platinum, taxanes have demonstrated activity in patients with GC, with docetaxel and paclitaxel both showing a survival benefit in first-line and second-line treatment of patients with metastatic GC, respectively.14 Recently, the use of docetaxel has shown survival improvements when combined with S1 and platinum-based therapy in the perioperative setting (FLOT4).4 The triplet regimen of S1/platinum/docetaxel is considered a standard of care in several Western nations.4 Similarly, in Japan, addition of docetaxel to S1 adjuvant chemotherapy has demonstrated a survival benefit.6 However, other GC trials investigating the role of taxanes in the adjuvant and metastatic setting were unable to demonstrate survival improvement.7 8 These contradictory results underpin the urgent clinical need for predictive biomarkers to identify patients who will benefit from taxane chemotherapy. While there are data for biomarkers predicting S1 and platinum benefit, there are currently no predictive biomarkers for taxane therapy in GC described in the literature.9–13

Stomach cancer Adjuvant Multi-Institutional group Trial (SAMIT) was a phase III randomised trial conducted in Japan, to assess superiority of adjuvant sequential treatment (paclitaxel followed by UFT or paclitaxel followed by S1) compared with monotherapy (UFT or S1 alone), and to assess the non-inferiority of UFT compared with S1 after curative surgery.14 UFT is an oral combination of uracil (an inhibitor of dihydropyrimidine dehydrogenase (DPD)) and tegafur (a S1 prodrug).15 S1 is an oral combination of tegafur, gimeracil (a DPD inhibitor) and oteracil (a compound that localises preferentially to the gut and inhibits orotate phosphoribosyl-transferase, reducing activation of S1 and gastrointestinal toxicity).16 The clinical results of SAMIT, reported in 2014, showed that sequential paclitaxel did not improve disease-free survival (DFS) and UFT was not non-inferior to S1.14

Taxanes are thought to exert their anti-cancer effects through aberrant stabilisation of microtubules, leading to defects in chromosome segregation, mitotic arrest and activation of the spindle assembly checkpoint, where prolonged activation results in cell death. Studies in patients with breast cancer suggest that altered expression of genes involved in the spindle assembly checkpoint may affect cellular sensitivity to paclitaxel,17 18 and immunogenic cell death after chemotherapy.19 20 We hypothesised that in GC the expression of genes involved in chromosomal stability and/or immunogenic cell death may predict benefit from paclitaxel. We aimed to test this hypothesis using samples and clinical data from the SAMIT trial. Findings from the SAMIT study were validated in an independent phase 2 GC trial cohort from South Korea.

METHODS
SAMIT trial
SAMIT was a 2×2 factorial design phase III study conducted in Japan.14 In total, 1495 patients were randomised to either single agent UFT (UFT arm), single agent S1 (S1 arm) or sequential paclitaxel-S1/Pac-S1 arm) or sequential paclitaxel-UFT (Pac-UFT arm).

Surgical sample processing and RNA extraction for biomarker cohort
Formalin-fixed, paraffin-embedded (FFPE) blocks or unstained cut sections from gastrectomy specimens were collected from Japanese sites participating in the SAMIT study.RNA was extracted from the primary tumour and prepared for NanoString profiling (online supplemental methods).

NanoString analysis
A custom designed NanoString (NanoString Technologies, USA) panel of 476 genes was used (online supplemental table 1). Gene ontologies were broadly categorised into:
1. Genes involved in the spindle assembly checkpoint.
2. Genes with therapeutic implications in GC.
3. Genes involved in the GC tumour microenvironment.
4. Genes relevant to oncogenic signalling pathways.
5. Genes with frequent genomic alterations in GC.
6. Immune-related genes in GC.
7. DNA damage repair genes.
8. Genes with reported predictive benefit in GC chemotherapy from the literature.

Quality control (QC) analyses between triplicates and corrections for batch effects are described in the supplement (online supplemental methods and figure S1). Normalised NanoString gene expression data and correlative clinicopathological characteristics are provided in online supplemental table 1.

Gene signature development using machine-learning models
The random forest method uses an ensemble of classification trees on several variables.21 Metrics for measuring the predictive performance included accuracy, precision, recall, F-measure and area under curve (AUC). The F-measure has been shown to handle class imbalances in the dataset better than positive predictive value and accuracy.22 23

Random forests return a prediction from a collection of classification trees (in our analyses we used ntree=3000) (online supplemental methods). Each tree is grown by using a bootstrap sample of the data set and, only a random subset of the original variables is examined at each node. Multiple iterations of the random forest algorithm were run, performing a variable importance analysis and eliminating variables that did not contribute to the classification resulting in improved predictive performance.24 The model was trained based on a DFS< or ≥2 years classifier. As the random forest algorithm works best when trained on equally sized groups, this clinically relevant cut-off of DFS was also selected to create relatively equal groups. A sensitivity and variable importance analysis for F-measure was performed to identify the optimal number of genes to be included in the gene signature. The classes identified by the gene signature derived from the random forest algorithm were labelled 'Pac-Sensitive' (paclitaxel-sensitive, ie, those patients that derive a survival benefit from paclitaxel) and 'Pac-Resistant' (paclitaxel-resistant, ie, those patients that lack a survival benefit from paclitaxel). The randomForest R package was used.

Creation of training and validation cohorts for machine learning
The aim was to train the random forest model on a training cohort to generate a classifier, which would then be applied on a validation cohort to confirm accuracy, F-measure and AUC of the classifier. The classifier would then be tested on an ‘external validation cohort’ (described in the next section). AUC was calculated using a time-dependent receiver operating characteristic (ROC) curve for 2-year DFS, using the risksetROC package.

We took advantage of the unique 2×2 factorial design of the SAMIT trial. Essentially, each arm of the study could function as individual cohort of patients for training and validation. As SAMIT was a randomised phase III trial, the cohorts (arms) would be well balanced with respect to confounding factors. Thus, two paclitaxel treated (Pac-S1 and Pac-UFT) and
two paclitaxel untreated (S-1 and UFT) cohorts were formed. Tumour samples from one paclitaxel treated arm would be used to train the model, while the other would be used to validate the classifier. Since S-1 currently represents the standard of care for GC chemotherapy in Japan, we elected to use the Pac-S-1 arm (n=128) as the training cohort and the Pac-UFT arm (n=123) as the validation cohort. We also tested the interaction between the classifier and treatment with paclitaxel by applying the classifier on the non-paclitaxel containing arms (S-1 and UFT).

External independent validation South Korean paclitaxel-ramucirumab trial cohort (Pac-Ram)

Data for this validation cohort were derived from a single-arm non-randomised phase II trial of patients with GC with metastatic disease treated with paclitaxel and ramucirumab (n=47).22 Patients with metastatic GC who had progressed on at least 1 line of chemotherapy which included platinum/fluoropyrimidine were enrolled in a phase II trial at the Samsung Medical Center, South Korea. All patients provided written informed consent before enrolment. A tumour biopsy was obtained between day −42 and day 1 prior to initiation of study treatment. Intravenous paclitaxel and ramucirumab were administered until documented disease progression, unacceptable toxicity or patients’ refusal was reported. RNA-Seq was performed and data were aligned to GENCODE V.19 transcript annotation using STAR. Transcripts per million abundance measure were generated using RSEM. RNA-seq transcripts mapping to genes profiled using the NanoString panel were extracted. Random forest classifiers trained on the SAMIT study were applied to the normalised RNA-seq data to predict benefit from paclitaxel therapy.

The Cancer Genome Atlas analysis

Gene expression data and clinical data for the The Cancer Genome Atlas (TCGA) stomach adenocarcinoma (STAD) cohort were downloaded from Firebrowse.26 Illumina HiSeq RNA-SeqV2 RSEM normalised gene values were used and applied through a similar pipeline as the Pac-Ram cohort RNA-Seq samples for generation of the random forest gene signature. HER2 status of TCGA STAD samples were derived from the HER2 index of Li et al.27 The HER2 index is an expression-based classifier reflecting the HER2-enriched transcriptional pattern for tumours harbouring HER2 aberrations. A cut-off of 0.75 was used to classify samples as positive.

Statistical analyses

The primary objective of this translational study was to identify a gene signature generated by a machine-learning model that predicts survival benefit from paclitaxel chemotherapy in patients from the SAMIT trial, and to validate these findings in an independent GC patient cohort.

All analyses were done using R (V.3.6.1) with statistical significance set at two-tailed p<0.05. Fisher’s exact test was used to evaluate associations with categorical variables. Kaplan-Meier curves with log-rank statistics were used to compare overall survival (OS), DFS in the SAMIT cohort and progression-free survival (PFS) in the Korean cohort. OS was calculated from the date of randomisation to the date of death by any cause, DFS was calculated from the date of randomisation to the date of first event (relapse of stomach cancer, death from any cause or occurrence of a second cancer), PFS was calculated from the date of first dose of paclitaxel/ramucirumab to the date of disease progression or death from any cause). The assumption of proportional hazards was checked using the scaled Schoenfeld residuals method and was supported. Cox regression was performed to present HRs and 95% confidence intervals (CI). Univariate and multivariate analysis for DFS was performed.

RESULTS

SAMIT biomarker cohort characteristics

RNA was extracted from 552 GC resection samples (37% of the whole SAMIT trial population) and profiled by a custom-made NanoString assay. 53 samples (10%) failed QC and were excluded from further analyses leaving a total of 499 GGs for final analyses (figure 1A). There were no statistically significant differences in patient characteristics between the original SAMIT trial population and subset of patients in the biomarker cohort (online supplemental table S2). Patient characteristics were well balanced between the arms in the biomarker cohort (table 1). Median age was 65 years (range 29–80 years) and 69% (n=343) were male. The median follow-up time was 60.5 months (IQR: 47.5–78.3 months), and there were no survival differences between the biomarker cohort and original SAMIT trial population (online supplemental figure S2). There were also no differences in the clinicopathologic characteristics of the samples that failed QC and the ones that did not (online supplemental table S3).

Development of a predictive gene signature of paclitaxel benefit using random forest analysis

NanoString profiles did not reveal any significant differences in gene expression between treatment arms, and none of the single gene expression levels were associated with DFS by Cox univariate regression after correction for multiplicity using the false discovery rate method (online supplemental table 1 and figure 2A).

To identify gene signatures for paclitaxel-benefit, we found that application of the machine-learning approach by randomly dividing SAMIT samples into training and validation cohort failed to deliver a gene signature that could be validated in the Pac-Ram external cohort (online supplemental methods). We, therefore, decided to take advantage of the unique 2×2 factorial design of the SAMIT trial, with each arm of the study representing an individual cohort. The cohorts were well balanced with respect to clinicopathologic factors (table 1). Using the Pac-S-1 arm (n=128) as the training cohort, we labelled patients as Pac-Sensitive if they derived a survival benefit from paclitaxel and as Pac-Resistant if they did not have a survival benefit from paclitaxel (DFS≥ or <2 years) (figure 1B).

From the Pac-S-1 cohort, random forest and variable importance analysis identified the top genes contributing to the model, on which the classifier was trained. A sensitivity analysis using different numbers of genes within the signature was performed, using F-measure and accuracy (online supplemental figure S3), to decide on the optimal number of genes within the signature. A 19-gene signature (random forest gene signature) trained on the Pac-S-1 cohort and tested on the Pac-UFT validation cohort (n=123) had the highest accuracy and F-measure, and was selected for further analysis. In the Pac-UFT arm, patients classified as Pac-Sensitive by the random forest gene signature had a significantly longer DFS compared with those classified as Pac-Resistant: 3-year DFS 66% (Pac-Sensitive) vs 40% (Pac-Resistant) (HR 0.44, 95%CI 0.25 to 0.76, logrank p=0.0029) (figure 2B). Accuracy of the random forest gene signature in the Pac-UFT validation cohort was 0.61, F-measure was 0.71 and AUC was 0.75 (95% CI 0.50 to 0.99).

Clinicopathological features of the Pac-Resistant patients in Pac-UFT revealed no significant differences when compared...
Figure 1  Flow chart of sample analysis from SAMIT trial A: In total, 1,495 patients were randomised after surgery in the SAMIT trial to four arms (UFT alone, S-1 alone, Pac-UFT and Pac-S-1). After assessment of formalin-fixed paraffin embedded blocks (FFPE) for tumour content, RNA was extracted and available for 552 samples which were profiled using the NanoString platform. After quality control postprofiling, 499 samples were included in the final analysis. (B) Samples from the Pac-S1 arm were selected to train the random forest machine-learning model. Samples were trained using a 2-year DFS to define Pac-Sensitive and Pac-Resistant groups. Using 476 genes in the custom-made NanoString panel, a variable importance analysis was performed, and the top 19 genes selected. The final model was retrained with these 19 genes to generate a random forest gene signature. This signature was then applied on the Pac-UFT arm as an internal validation cohort and the S-1 and UFT arms to test for interaction with paclitaxel treatment. The Pac-Ram samples were tested as an external validation cohort. SAMIT, Stomach cancer Adjuvant Multi-Institutional group Trial.
with Pac-Sensitive patients (online supplemental table S4). When clinical features of the Pac-UFT cohort were considered, in both univariate and multivariate survival (DFS) analysis, depth of invasion (pT), lymph node status (pN) and random forest gene signature were significant predictors of survival (table 2). After adjusting the survival model for depth of invasion and lymph node status, the random forest gene signature remained significantly associated with survival (HR 0.45, 95%CI 0.26 to 0.80, p=0.006).

Validation of the random forest gene signature in an independent external cohort
After successful internal validation of the random forest gene signature, we tested the signature in an independent external cohort (Pac-Ram cohort). In the Pac-Ram cohort, the random forest gene signature predicted that patients with Pac-Sensitive GC have longer PFS compared with patients with Pac-Resistant GC (median PFS 147 days vs 112 days, HR 0.48, 95%CI 0.25 to 0.91, logrank p=0.022) (figure 2C). In the Pac-Ram cohort, F-measure was 0.62, accuracy 0.64 and AUC was 0.88 (95% CI 0.68 to 1.0). This suggests that the random forest model trained on the Pac-S1 cohort identified a robust classifier which was able to predict paclitaxel benefit in both, the internal and the external validation cohorts. There was no association between the random forest gene signature and radiologically measured objective response rates in the Pac-Ram cohort.

To study the interaction of the random forest gene signature with paclitaxel treatment, we ran the model on the UFT alone and S-1 alone arms of the SAMIT study and compared results to the Pac-UFT arm. There was no survival difference between patients with Pac-Sensitive and Pac-Resistant GC in the UFT or S-1 arms (HR 0.99, 95%CI 0.55 to 1.83, p=0.99)(supplementary figure S4). The test of interaction for the paclitaxel containing regimens with random forest gene signature was significant (p<0.001). The test for heterogeneity to validate the interaction of paclitaxel on DFS was significant (p<0.001). These findings suggest that the random-forest gene signature is predictive for survival benefit with paclitaxel treatment and is not merely a prognostic biomarker.

Transcriptomic characteristics of GC classified by the random forest gene signature
To understand the expression profiles of individuals within the random forest gene signature, the 19 selected genes were studied in greater detail in the entire SAMIT dataset. The 19 genes in the signature include: CD209, TOP3B, BCL2, MCM2, RAD9A, IL10, HLA-DRB1, NME1, FAP, FANCM, ADORA2A, MS4A1, FANCG, CD3G, DSCR6, TBX21, ZWILCH, IL17A and FCGR3A. Heatmaps of genes split by Pac-Sensitive and Pac-Resistant GC revealed some differences in gene-expression profiles, although these are unlikely to have led to separate clusters using traditional unsupervised hierarchical clustering techniques (online supplemental figure 5A). Pairwise comparisons of genes revealed that some genes were significantly higher expressed in Pac-Sensitive GC compared with Pac-Resistant GC (or vice-versa), while others had similar median expression levels between the two groups (online supplemental figure 5B). These findings highlight the importance of considering multigene interactions for the generation of a signature, which may be achieved through machine-learning pattern recognition.

We tested the random forest gene signature on TCGA STAD samples (n=375), and 76% were classified as Pac-Sensitive. Of the four TCGA subtypes: chromosomal instable (CIN), genome stable (GS), microsatellite instable and Epstein-Barr virus associated, 29% GS GC were more frequently classified as Pac-Sensitive than Pac-Resistant (25% vs 9%), while CIN GC were more frequently classified as Pac-Sensitive (64% vs 51%); Fisher’s exact p=0.0019 (figure 3A). Mutation counts were higher in the Pac-Sensitive group (Wilcoxon p=0.00031) (figure 3B). There were no differences in the HER2 status between Pac-Sensitive and Pac-Resistant GC. (Fisher’s exact p=0.68) (figure 3C).

Finally, the association between the Pac-Resistant/Pac-Sensitive groups and immune-related gene expression profiles was explored. The customised NanoString panel used in this study included 79 immune-related genes. In the SAMIT samples, Pac-Sensitive GC had a higher expression of markers of cytolytic T-cells such as CD8A and PRF1 (figure 3D). The expression

---

**Table 1** Patient characteristics of biomarker cohort of SAMIT study

<table>
<thead>
<tr>
<th>Factors</th>
<th>UFT (n=124)</th>
<th>S-1 (n=124)</th>
<th>Pac-UFT (n=123)</th>
<th>Pac-S-1 (n=128)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Median (range)</td>
<td>65 (29–80)</td>
<td>65 (38–80)</td>
<td>66 (30–80)</td>
<td>67(37–80)</td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
<td>87</td>
<td>77</td>
<td>86</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>37</td>
<td>47</td>
<td>37</td>
<td>35</td>
</tr>
<tr>
<td>Performance status</td>
<td>0</td>
<td>107</td>
<td>103</td>
<td>101</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>17</td>
<td>21</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>Histological type</td>
<td>Differentiated</td>
<td>45</td>
<td>52</td>
<td>49</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Undifferentiated</td>
<td>77</td>
<td>70</td>
<td>71</td>
<td>72</td>
</tr>
<tr>
<td>T stage</td>
<td>pT1</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>pT2</td>
<td>40</td>
<td>38</td>
<td>32</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>pT3</td>
<td>77</td>
<td>80</td>
<td>83</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>pT4</td>
<td>5</td>
<td>2</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>N stage</td>
<td>pN0</td>
<td>17</td>
<td>25</td>
<td>34</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>pN1</td>
<td>46</td>
<td>51</td>
<td>41</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>pN2</td>
<td>59</td>
<td>46</td>
<td>46</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>pN3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

---

American Joint Committee on Cancer/Union for International Cancer Control (AJCC/UICC) sixth edition.
Figure 2  Performance of random forest gene signature. (A) Heatmap of NanoString expression profiles from SAMIT samples by treatment arm. NanoString gene expression is represented in columns, scaled. Blue to red denotes transcript expression, with blue indicating low gene expression and red indicating high gene expression. There is no global difference in gene expression profiles between the four arms. (B) Kaplan-Meier curve of disease free survival (DFS) of patients classified by the random forest gene signature as either Pac-Sensitive (blue) or Pac-Resistant (red) in the validation Pac-UFT arm. A 3-year DFS 66% (Pac-Sensitive) vs 40% (Pac-Resistant) (HR 0.44, 95% CI 0.25 to 0.76, logrank p=0.0029) (C) Kaplan-Meier curve of progression-free survival (PFS) of patients classified by the random forest gene signature as either Pac-Sensitive (blue) or Pac-Resistant (red) in the external validation Pac-Ram cohort. Median PFS 147 days vs 112 days, HR 0.48, 95% CI 0.25 to 0.91, logrank p=0.022. SAMIT, Stomach cancer Adjuvant Multi-Institutional group Trial.
of immune checkpoints genes such as CTLA-4, PD-1 and LAG3 was also higher in Pac-Sensitive GC. Similar associations with the expression of immune-related genes were also observed in the Pac-Ram and TCGA datasets (figure 3E,F).

DISCUSSION

The 2×2 factorial trial design of the SAMIT study and the relatively large number of collected tumours provided a good experimental setting to search for biomarkers of paclitaxel benefit in patients with resectable GC. The SAMIT trial design allowed the creation of training and validation cohorts treated in a uniform manner with balanced patient and tumour characteristics between treatment arms. The presence of paclitaxel and non-paclitaxel containing treatment arms allowed us to distinguish between the predictive and prognostic value of potential biomarkers, and to identify treatment interactions specific to paclitaxel. Given the limited amount of FFPE tissue collected from the trial patients, we chose to create a customised NanoString panel to investigate the expression of a relatively large number of genes to identify biomarkers of paclitaxel benefit. The NanoString platform has proven to provide reliable results when performed on RNA extracted from FFPE tissue. Previous studies have shown good correlation between NanoString and RNAseq results in GC.30,31 Based on these data, we used RNA-seq data from a separate trial of GC patients treated with paclitaxel (and ramucirumab) as external validation.

A comprehensive literature review was performed to select the probes for the NanoString panel, covering a range of mechanisms of action of paclitaxel and cell death, GC-specific oncogenesis and other previously proposed predictive markers of taxane benefit. We chose a machine-learning model to identify a gene signature that was potentially predictive of paclitaxel benefit. The random forest method has been used by other groups as a reliable machine-learning algorithm for gene classifier creation.12,32 One advantage of the random forest method is the ability to incorporate a large number of variables (in this situation, genes) in the model. Few algorithms can process data with a variable size which is much larger than the number of samples in the training dataset and continue to demonstrate significant predictive performance.34

Recently, a four gene classifier was described to predict benefit of adjuvant chemotherapy in GC.34 This classifier was the first to suggest that GC patients classified as ‘immune high’ and low-risk derive no benefit from adjuvant cytotoxic chemotherapy after curative D2 gastrectomy. It is interesting to note that the gene signature identified in our study includes immune related genes such as CD209, a dendritic-cell marker, T-cell surface glycoprotein gamma chain CD3G and interleukin genes IL10 and IL17A.35,36 The Pac-Resistant and Pac-Sensitive GC also appeared to have different expression of other immune-related genes including immune checkpoints and markers of T-cell cytolytic activity.29,31 Other studies have suggested that patients with immune cell-rich colorectal cancers do not benefit from chemotherapy.37 Our random forest gene signature included cell-cycle checkpoint gene RAD9A and BCL2, a regulator of apoptosis.8,18,39 RAD9A is a component of the 9-1-1 DNA clamp, which has multiple roles in DNA repair. As Pac-Sensitive tumours had higher levels of RAD9A, it could be possible that paclitaxel induced DNA damage in the setting of high RAD9A expression increases mitotic cell death. RAD9A has also been reported to bind and neutralise BCL2, thereby inducing a proapoptotic state.40 Targeting the BCL2 pathway has been shown to reverse paclitaxel resistance in oesophageal cancer cell lines.41 TOP3B, a DNA topoisomerase enzyme included in our gene signature, was found to be prognostic when overexpressed in ovarian cancer.42 It is likely that there are interactions, yet to be fully defined, between several genes in these important pathways which subsequently lead to sensitivity or resistance to paclitaxel.

An important point of consideration is the similarities and differences between the SAMIT trial patients and the external Pac-Ram cohort. Both cohorts included GC patients treated with paclitaxel. However, in SAMIT, paclitaxel was combined with 5FU based chemotherapy, while in the Pac-Ram cohort paclitaxel was combined with ramucirumab, an anti-angiogenic monoclonal antibody. SAMIT was a study of patients treated in the adjuvant setting including earlier stage tumours, while Pac-Ram comprised patients with late stage GC treated in the metastatic setting after progression on first-line platinum-based chemotherapy. While DFS was used to measure survival in the SAMIT study, PFS was measured in the Pac-Ram clinical trial. Despite these differences, the random forest gene signature was able to predict paclitaxel benefit in both cohorts. The supporting data from the 5FU arms suggests that the signature is not a prognostic biomarker. These findings suggest that this signature is relatively specific for predicting benefit from paclitaxel chemotherapy in GC. Our ability to analyse RNA from only a subset of the original trial population is a limitation of this study. However, clinicopathological variables and survival were similar between our biomarker study cohort and the original cohort. As the paclitaxel-containing cohorts in our study were combined with various other agents (S-1, UFT and ramucirumab), the effects of these agents on the performance of the random forest

Table 2  Univariate and multivariate survival analysis in Pac-UFT

<table>
<thead>
<tr>
<th>Factors</th>
<th>Univariate</th>
<th>Multivariate</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=123</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>P value</td>
</tr>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>P value</td>
</tr>
<tr>
<td>Random forest gene signature</td>
<td>0.44 (0.25 to 0.76)</td>
<td>0.0038</td>
</tr>
<tr>
<td>Pac-Sensitive versus Pac-Resistant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>1.00 (0.97 to 1.03)</td>
<td>0.79</td>
</tr>
<tr>
<td>Gender: Female versus male</td>
<td>0.70 (0.38 to 1.28)</td>
<td>0.25</td>
</tr>
<tr>
<td>ECOG Performance Status: PS 1 versus PS 0</td>
<td>1.27 (0.66 to 2.47)</td>
<td>0.47</td>
</tr>
<tr>
<td>Histology</td>
<td>1.43 (0.81 to 2.51)</td>
<td>0.22</td>
</tr>
<tr>
<td>Undifferentiated/others versus differentiated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth of invasion pT3/4 vs pT1/2</td>
<td>2.48 (1.17 to 5.26)</td>
<td>0.018</td>
</tr>
<tr>
<td>Lymph node status pN2/3 vs pN0/1</td>
<td>2.89 (1.69 to 4.94)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

AJCC/UICC sixth edition. Significant p values (< 0.05) in bold.
Figure 3  Transcriptomic characteristics of random forest gene signature. (A) Alluvial plot of TCGA STAD samples by gastric cancer (GC) subtype. TCGA samples (n=375) were divided into Pac-Sensitive and Pac-Resistant by the random forest gene signature. These groups were correlated with TCGA GC subtypes chromosomal instable (CIN), genome stable (GS), microsatellite instable (MSI) and Epstein-Barr virus associated (EBV). There was a statistically higher proportion of GS patients in the Pac-Resistant group (Fisher’s exact p=0.0019). (B) Violin plot of mutation count between Pac-Sensitive and Pac-Resistant TCGA samples. Mutation count was higher in the Pac-Sensitive group (Wilcoxon p=0.00031). (C) Alluvial plot of TCGA STAD samples by HER2 status. Samples were correlated with HER2 status, with no significant difference. (D) Volcano plot comparing immune-related gene expression between Pac-Sensitive and Pac-Resistant in SAMIT. X-axis: log2 fold change (log2FC) of gene expression between Pac-Sensitive and Pac-Resistant. Y-axis: log10 adjusted p values after false discovery rate correction. Genes of interest have been annotated within the plot. Grey dots represent genes with similar expression in Pac-Sensitive and Pac-Resistant GC. Blue dots represent genes which are overexpressed in Pac-Sensitive GC. Red dots represent genes which are overexpressed in Pac-Resistant GC. (E) Volcano plot of immune-related genes comparing gene expression level between Pac-Sensitive and Pac-Resistant in Pac-Ram cohort. (F) Volcano plot of immune-related genes comparing gene expression level between Pac-Sensitive and Pac-Resistant in TCGA STAD samples. SAMIT, Stomach cancer Adjuvant Multi-Institutional group Trial; STAD, stomach adenocarcinoma; TCGA, The Cancer Genome Atlas.
gene signature could not be fully determined due to the lack of a GC patient cohort treated with paclitaxel alone. The accuracy and F-measure in our internal validation cohort was moderate (0.61 and 0.71, respectively). The signature, however, was validated in an external cohort, and further study is required to improve these metrics. As a precedent, similar measurements have been seen in more robust and validated tests such as Onco-type DX in breast cancer, a tool to predict for lack of benefit for chemotherapy in hormone receptor positive breast cancer.45 In the initial oncotype dx report performed on National Surgical Adjuvant Breast and Bowel Project B-14 and B-20 studies, accuracy was 72% and F-measure was 76% (calculated from data provided in manuscript), which is similar to our study. Another limitation of this study is the post hoc nature of this translational analysis. A prospective clinical trial would be required to confirm the clinical utility of this gene signature.

In conclusion, this study is the first and largest performed to identify a gene signature predictive of paclitaxel benefit in GC. Given the increasing use of taxanes in GC in the adjuvant and perioperative setting, this biomarker has significant potential to guide clinicians in identifying patients with GC who might benefit from taxane based therapy. Validation of the gene signature in prospective trials is warranted.

Author affiliations
1 Department of Haematology-Oncology, National University Cancer Institute Singapore, National University Hospital, Singapore
2 Yong Loo Lin School of Medicine, National University of Singapore, Singapore
3 Program in Cancer and Stem Cell Biology, Duke-NUS Medical School, Singapore
4 The N.I Institute for Health, National University of Singapore, Singapore
5 Biostatistics Unit, Yong Loo Lin School of Medicine, National University Singapore, Singapore
6 Department of Surgical Oncology, Gifu University Graduate School of Medicine, Gifu, Japan
7 Department of Gastric Surgery, National Cancer Center Hospital, Tokyo, Japan
8 Kanagawa Cancer Center Research Institute, Yokohama, Japan
9 Department of Surgery, Yokohama City University, Yokohama, Japan
10 Department of Clinical Biostatistics, Graduate School of Medicine, Kyoto University, Kyoto, Japan
11 Tokai Central Hospital, Kamakigahara, Japan
12 Cancer Science Institute of Singapore, National University of Singapore, Singapore
13 Laboratory of Cancer Epigenome, Department of Medical Sciences, National Cancer Centre Singapore, Singapore
14 InSilico Genomics, Phoenix, Arizona, USA
15 Department of Pathology, GROW - School for Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht, The Netherlands
16 Division of Pathology and Data Analytics, Leeds Institute of Medical Research at St James’s, University of Leeds, Leeds, UK
17 Department of Medicine, Division of Hematology-Oncology, Samsung Medical Center, Gangnam-gu, Republic of Korea
18 Department of Gastrointestinal Surgery, Kanagawa Cancer Center, Yokohama, Japan
19 Department of Surgery, Otsuka Hospital, Odawara, Japan
20 Genome Institute of Singapore, Singapore
21 SingHealth/Duke-NUS Institute of Precision Medicine, National Heart Centre Singapore, Singapore
22 Department of Physiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

Correction notice This article has been corrected since it published Online First. Online supplementary table 1 has been added.

Collaborators We are deeply grateful to Kazuaki Tanabe (Hiroshima University), Michiya Kobayashi (Kochi University), Shigeumi Yoshino (Yamaguchi University), Masazumi Takahashi (Yokohama Citizens Hospital), Nobuhiro Takiguchi (Chiba Cancer Center), Norio Mitsumori (Tokyo Jikei University), Kazuma Fujitani (Osaka Prefectural General Medical Center), Ryoji Fukushima (Teikyo University), Isao Noguchi (Shikoku Cancer Center), Yoshihiro Kakechi (Kobe University), Naoki Hibiayashi (Hiroshima City Asa Citizens Hospital), Yukihiko Tokunaga (Osaka Kita Teishin Hospital), Akioi Takagane (Hokodate Goryokaku Hospital), and Kazuhiro Nishikawa (Osaka Medical Center) for providing clinical samples from the SAMIT trial.
Supplemental material

This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access

This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID IDs

Heike I. Grabsch http://orcid.org/0000-0001-9520-6228

Patrick Tan http://orcid.org/0000-0002-0179-8048

REFERENCES


