Practical biomarkers and robust multiplex models for the prediction of response to the promising first-line chemotherapy: A theranostic study in metastatic ovarian cancer patients with residual peritoneal tumors

Reika Kawabata-Iwakawa
Gunma University Initiative for Advanced Research, Gunma University

Norihiro Iwasa
Saitama Medical University International Medical Center

Kenichi Satoh
Shiga University

Jacques Colinge
Institute of Cancer Research of Montpellier (IRCM), Inserm, University of Montpellier, ICM

Muneaki Shimada
Tohoku University Graduate School of Medicine

Satoshi Takeuchi
Kobe Tokushukai Hospital

Hiroyuki Fujiwara
Jichi Medical University

Hidetaka Eguchi
Juntendo University Graduate School of Medicine

Tetsuro Oishi
Tottori University School of Medicine

Toru Sugiyama
Iwate Medical University

Mitsuaki Suzuki
Jichi Medical University

Kosei Hasegawa
Saitama Medical University International Medical Center

Keiichi Fujiwara
Saitama Medical University International Medical Center

Masahiko Nishiyama (✉ m.nishiyama@gunma-u.ac.jp )
Gunma University

Research Article

Keywords: Predictive biomarker, Personalized medicine, Ovarian cancer, Intraperitoneal carboplatin plus intravenous dose-dense paclitaxel (ddTCip) therapy, The first-line chemotherapy

Posted Date: January 19th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2490080/v1

License: ©  This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Background: In advanced or metastatic ovarian cancer patients, the therapeutic impact of molecular targeted agents and immunotherapy is limited, and current chemotherapeutic algorithm is still far from personalized medicine. We recently demonstrated that intraperitoneal carboplatin with dose-dense paclitaxel (ddTCip) therapy is a promising front-line chemotherapy even in the patients with residual peritoneal tumors, which led us to this theranostic study for biomarker discovery to realize the precision medicine (ID: UMIN000001713 on Feb 16 th , 2009).

Methods: We first validated previously suggested markers (41 genes and 3 predictive models for the therapeutic efficacy and 31 polymorphisms for the toxicity), sought out more active effective biomarkers through genome-wide transcriptome and genotyping analyses, and then developed multiplex statistical prediction models for progression free-survival (PFS) and toxicity. Multiple regression analysis following forward stepwise method and Classification and Regression Trees (CART) algorithm were mainly employed to develop multiplex prediction models.

Results: The association analyses with PFS in 76 patients followed by the validation study using data sets in 189 patients published in The Cancer Genome Atlas revealed that SPINK1 expression could be a possible predictive biomarker of ddTCip efficacy even when used alone, and multiple regression analyses provided a potent efficacy prediction model using expression data of 5 genes. SPINK1 appeared to be a critical resistant determinant of ddTCip therapy, which indicates the potential of SPINK1 also to be a novel therapeutic target. As for the toxicity prediction, ABCB1rs1045642 and ERCC1rs11615 polymorphisms appeared to closely associate with grade2-4 hematologic toxicity and peripheral neuropathy, respectively. We further successfully composed robust multiplex prediction models for the adverse events-CART models using a total of 4 genotype combinations and further powerful multiple regression models using 15 polymorphisms on 12 genes.

Conclusions: We newly proposed SPINK1 expression as a powerful predictive biomarker of the efficacy for ddTCip therapy and confirmed the predictive values of ABCB1 and/or ERCC1 polymorphisms for the toxicity. Multiplex prediction models composed herein were also found to work well for the prediction of therapeutic response. These may raise the potential to realize a precision medicine in the essential treatment for metastatic ovarian cancer patients.

Introduction

Ovarian cancer is often asymptomatic, leading to delays in diagnosis. At the diagnosis, the cancer is advanced with peritoneal dissemination in many cases [1, 2]. The most common treatment to these patients remains to be debulking surgery with pre- and post-operative adjuvant platinum-based chemotherapy, largely combined with taxanes [1–3]. Generally, the chemotherapy is initially active, but the consequent recurrence or relapse occurs in many patients. The response varies among patients, and not a few patients undergo a regimen without any obvious treatment benefit. However, despite of extensive efforts, any definitive predictive biomarkers of platinum/paclitaxel chemotherapy remain to be undetermined [4–19]. Further improvements in the outcome are dependent upon development of a more active therapeutic modality of the combination and an identification of the disease-specific molecules or genomic alterations that can be used as predictive biomarkers of the therapy.

We recently examined the novel modality, intraperitoneal (IP) carboplatin in combination with dose-dense paclitaxel (ddTCip), in advanced ovarian cancer and demonstrated that front-line chemotherapy with ddTCip was effective even for patients with suboptimal residual ovarian cancer [20]. In parallel with the phase II study, we performed this theranostic study to identify truly active predictive biomarkers of the efficacy and the toxicity induced by the therapy. Here, we first examined previously reported biomarkers (41 single marker genes and 3 multiplex predictive models for the therapeutic efficacy, and 31 polymorphisms for the toxicity), and then sought out more powerful markers and developed multiplex prediction models through genome-wide screening study in 76 patients and integrated trans-OMICS data analyses using a large-scale public database.

Materials And Methods

Patients and tissues

Between March 2009 and March 2012, a total of 117 patients with epithelial ovarian cancer (EOC) or primary peritoneal cancer, FIGO stage II–IV, who were considered eligible for this study, were temporarily registered prior to primary surgery. The patients underwent primary surgery, and if the residual tumor was judged to be measurable by the surgeon, IP port for carboplatin administration was placed during the operation. Of the 117 patients, 76 met the inclusion criteria and were enrolled. All patients had macroscopically residual peritoneal tumors after the initial debulking surgery, with ECOG Performance Status scores of 0–2, and had not undergone any chemotherapy or radiation therapy. All patients received ddTCip therapy: Weekly intravenous chemotherapy, largely combined with taxanes [1–3]. Generally, the chemotherapy is initially active, but the consequent recurrence or relapse occurs in many patients. The response varies among patients, and not a few patients undergo a regimen without any obvious treatment benefit. However, despite of extensive efforts, any definitive predictive biomarkers of platinum/paclitaxel chemotherapy remain to be undetermined [4–19]. Further improvements in the outcome are dependent upon development of a more active therapeutic modality of the combination and an identification of the disease-specific molecules or genomic alterations that can be used as predictive biomarkers of the therapy.

Fresh tumor specimens and peripheral blood samples were collected from all the patients. Tumor specimens were collected at surgery immediately frozen and respectively stored at -80°C. Blood samples were separated to the components by centrifugation and stored at -20°C until use. Ethics Committees at Gunma University, Saitama Medical University, Tottori University School of Medicine, Iwate Medical University and Jichi Medical University approved this research protocol, which was registered at UMIN Clinical Trials Registry (ID: UMIN000001713) on Feb 16th, 2009. All the patients provided written informed consent including donation of their peripheral blood and surgical specimens before inclusion in the study.

Extraction and purification of total RNA

Page 2/15
Total RNA was extracted from ovarian cancer tissues using RNA Nucleospin RNA II kit (Macherey–Nagel, Düren, Germany) according to the manufacturer’s protocols. For microarray analysis, the quality of total RNA was examined using RNA 6000 Nano Kit (Agilent Technologies, Santa Clara, U.S.) and Bioanalyzer 2100 (Agilent Technologies). The 2100 Bioanalyzer Expert software program was used to assign an RNA integrity number (RIN). Only RNA samples showing a RIN score equal to or greater than 7.1 were used for the further analyses.

**Real-time RT-PCR (reverse transcription polymerase chain reaction)**

One µg of total RNA was converted to cDNA using ReverTra Ace (Toyobo, Osaka, Japan) with random primer (9 mer) according to the manufacturer’s instructions. Primer and probe set for each gene other than ACTB (actin beta gene) were designed using The Probe Finder software in the Universal Probe Library (UPL) Assay Design Center (Roche Applied Science, Mannheim, Germany). Information of the sequences of the primers and UPL probes (Roche) was listed in Supplementary Table S1. For ACTB gene, Pre-Developed TaqMan Assay Reagents were purchased from Applied Biosystems (Applied Biosystems, Waltham, MA, U.S.). Each reaction was carried out in triplicate using qPCR QuickGoldStar Mastermix Plus reagent (Eurogentec, Liége, Belgium) and LightCycler 480 II system (Roche). Serial dilution sets of the positive controls were used for the quantification. Finally, the relative expression levels of each gene were calculated as a ratio to geometric mean of the ACTB expression levels.

**DNA microarray analysis**

Quality-checked total RNAs (0.5 µg) being sufficiently subjected to DNA microarray analysis were obtained from 55 patients. The samples were reverse transcribed to first-strand cDNA using MMLV (Moloney Murine Leukemia Virus) reverse transcriptase and T7 primer, and the Cy3-labelled cRNA were generated by the Quick Amp Labeling Kit One-Color (Agilent Technologies, Santa Clara, CA, U.S.). The labelled cRNA was purified by RNasy Kit (Qiagen, Valencia, CA, U.S.), and a total of 1.65 µg of cRNA was hybridized to Whole Human Genome Oligo 4x 44K (Agilent Technologies). Hybridization and washing procedure were conducted according to the manufacturer’s recommendation. Finally, the microarrays were scanned using an Agilent DNA Microarray Scanner (Agilent Technologies) and analyzed with Agilent Feature Extraction software version 9.5. Expression levels were normalized to the 75th percentile expression value of the entire spot using GeneSpring GX (Agilent Technologies).

**DNA extraction and genotyping**

Genomic DNA was extracted from a stored buffy coat from peripheral blood (7 mL) using NucleoSpin Tissue kit (Macherey-Nagel, Düren, Germany) following the manufacturer’s instruction. Single nucleotide polymorphisms (SNPs) in ABCB1 (ATP Binding Cassette Subfamily B Member 1 gene), ABCC1 (ATP Binding Cassette Subfamily C Member 1 gene), ABCC2 (ATP Binding Cassette Subfamily C Member 2 gene), CYP1B1 (Cytochrome P450 Family 1 Subfamily B Member 1 gene), CYP2C8 (Cytochrome P450 Family 2 Subfamily C Member 8 gene), CYP3A4 (Cytochrome P450 Family 3 Subfamily A Member 4 gene), CYP3A5 (Cytochrome P450 Family 3 Subfamily A Member 5 gene), ERCC1 (Excision Repair Cross-Complementation Group 1 gene), ERCC2 (Excision Repair Cross-Complementation Group 2 gene), GSTP1 (Glutathione S-Transferase Pi 1 gene), UGT1A1 (UDP Glucuronosyltransferase 1 family, polypeptide A1 gene), XRCC1 (X-Ray Repair Cross Complementing 1 gene), and XRCC3 (X-Ray Repair Cross Complementing 3 gene) were determined using TaqMan® Drug Metabolism Genotyping Assays or TaqMan® SNP Genotyping Assays (Life Technologies Co., Carlsbad, CA, USA), LightCycler® 480 Probes Master (Roche Applied Science, Indianapolis, IN, USA), and LightCycler® 480 System (Roche). GSTM1 (Glutathione S-transferase M1 gene) and GSTT1 (Glutathione S-Transferase Theta 1 gene) null polymorphisms were determined by PCR amplification followed by 2% agarose gel electrophoresis as essentially described by Medeiros R. et al. [4, 16]. The primer sequences used for this study are listed in Supplementary Table S2.

**Statistical Analysis**

PFS (progression-free survival, day) was used as an efficacy indicator in this study (n = 76). PFS was defined as the time interval between registration and progression or death, whichever occurred first, or the last follow-up for patients alive without progression. Linear regression analyses between z-scored gene expression value and logarithmic transformed PFS were adopted to seek the potent predictive markers using datasets of 64 cases with fixed PFS (non-censored cases) in this prospective study (ddTCip cohort) and 189 cases published in The Cancer Genome Atlas (TCGA, https://gdc.cancer.gov/about-data/publications/ov_2011) (TCGA cohort) [10]. PFS curves were estimated using Kaplan-Meier method, while Cox proportional hazards model was used for the survival analysis. These survival analyses were carried out using datasets of all patients including censored and non-censored cases in both cohorts (ddTCip cases, n = 76; TCGA cases, n = 189). To construct a prediction formula using multiple gene expression data of individual patient, multiple regression analysis following forward stepwise method was performed as follows: 1) Removal of the outlier with > 3 SD (standard deviation) by robust estimation from the data of all patients enrolled in this prospective study (ddTCip cases), 2) Multiple regression analysis using screened datasets of ddTCip cases, 3) Estimation of the consistency between fitted value and observed value calculated by constructed formula using datasets of ddTCip cases, 4) Selection of the variables by forward stepwise regression using datasets of ddTCip cases. 5) Validation of the predictive accuracy of constructed formula using datasets of TCGA cases.

Toxicities were analyzed in 76 patients who received at least one dose of chemotherapy. Among 2 patients discontinued ddTCip treatment due to acute hypersensitivity reaction to paclitaxel and 1 patient was not assessable in the genotyping, thereby pharmacogenomic analyses being performed in a total of 73 patients. Adverse effects were determined using National Cancer Institute Common Terminology Criteria for Adverse Effects (CTCAE-NCI), version 3.0 as with the clinical phase II study. Due to the limited number of cases, anorexia, nausea, vomiting, constipation, and diarrhea were combined into digestive symptom and subjected to the construction of the prediction model.

Fisher’s exact test was performed to validate the association of single genotype with toxicity. We applied the Classification and Regression Trees (CART) algorithm using the R packages of r part (version 4.1.1) and party kit (version 1.2–16). Multiple logistic regression analysis was also performed to construct the prediction model for the toxicity through the following steps: 1) Logistic regression analysis using data sets when the number of event
occurrence cases was \( \geq 10 \), 2) Evaluation by using the Akaike information criterion (AIC) with adding one variable at a time. When the AIC was improved, the variable was incorporated into the prediction model. 3) Calculation of The VIF (Variance Inflation Factor) for collinearity of the variables by using rms R package (version 6.2-0). When the VIF was more than 10, the variable was removed from the prediction model. 4) Evaluation of the odds ratio of the alternative allele to the reference allele for each variable, the AUC (area under the curve) by using pROC R package (version 1.18.0), and positive discrimination rate (PDF) of the obtained model to estimate the predictivity of the adverse events.

All statistical analyses were conducted using an R software version 3.3.2. Functional enrichment analysis was done by Ingenuity Pathway Analysis (IPA version 2.3; QIAGEN, Germany).

**Results**

**Validation of previously suggested single markers in the prediction of therapeutic efficacy**

We first examined single prediction markers for the efficacy of platinum and/or paclitaxel chemotherapy previously suggested. Among a variety of exploratory molecular markers, we selected a total of 41 genes, whose clinical impact and the functional significance as a drug sensitivity determinant has been demonstrated in two or more reports on the National Library of Medicine's PubMed (https://pubmed.ncbi.nlm.nih.gov/) and/or in our previous studies [5, 21]. In 64 patients with fixed PFS (non-censored cases), these candidate markers were subjected to real-time RT-PCR and then analyzed for correlation of the expression levels with PFS, by linear regression analyses. These analyses in ddTCip cohort and the validation study using 189 data sets published in a large-scale TCGA database (TCGA cohort) indicated 4 highly correlative genes: They were **SPINK1** (Serine Peptidase Inhibitor Kazal Type 1 gene), **TNNT3** (Troponin T3, Fast Skeletal Type gene), **IRF9** (Interferon Regulatory Factor 9 gene) and **ABCC2** (ATP Binding Cassette Subfamily C Member 2 gene) (Table 1, Fig. 1). Nevertheless, for **IRF9**, the observed correlation slopes with PFS were directly opposite in the 2 cohort cases.
Table 1

Linear regression analysis: Correlation of the expression with progression free survival (PFS) in 41 genes previously reported as potent efficacy markers of taxane and/or platinum therapy. *ddTCip cohort, ovarian cancer patients who were enrolled in this theranostic study and received intraperitoneal carboplatin plus intravenous dose-dense paclitaxel; **TCGA cohort, published data sets in The Cancer Genome Atlas (ovarian cancer patients who received taxane and/or platinum therapy); ***Adj_R2, adjusted R (regression coefficient)-squared; **** p, p-value for the slope; #N/A, not analyzed

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
<th>ddTCip cohort*</th>
<th>TCGA cohort**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>non-censored 64 cases</td>
<td>Adj_R2****</td>
<td>slope</td>
</tr>
<tr>
<td></td>
<td>non-censored 189 cases</td>
<td>Adj_R2</td>
<td>slope</td>
</tr>
<tr>
<td>JOSD2</td>
<td>Josephin domain containing 2</td>
<td>0.081</td>
<td>-2.109</td>
</tr>
<tr>
<td>TUBB3</td>
<td>tubulin beta 3 class III</td>
<td>0.063</td>
<td>-7.383</td>
</tr>
<tr>
<td>PPFIA1</td>
<td>PTPRF interacting protein alpha 1</td>
<td>0.074</td>
<td>-8.436</td>
</tr>
<tr>
<td>IL6</td>
<td>interleukin 6</td>
<td>0.080</td>
<td>-2.846</td>
</tr>
<tr>
<td>IRF9</td>
<td>interferon regulatory factor 9</td>
<td>0.074</td>
<td>-8.436</td>
</tr>
<tr>
<td>TNNT3</td>
<td>troponin T3, fast skeletal type</td>
<td>0.085</td>
<td>-23.327</td>
</tr>
<tr>
<td>FXYO1</td>
<td>forkhead box 01</td>
<td>0.035</td>
<td>-2.109</td>
</tr>
<tr>
<td>ERCC2</td>
<td>ERCC excision repair 2, TFIIH core complex helicase subunit</td>
<td>0.017</td>
<td>-6.609</td>
</tr>
<tr>
<td>VEGF</td>
<td>vascular endothelial growth factor A</td>
<td>0.009</td>
<td>4.823</td>
</tr>
<tr>
<td>NDE1</td>
<td>nudE neurodevelopment protein 1</td>
<td>0.016</td>
<td>-1.925</td>
</tr>
<tr>
<td>MTA1</td>
<td>metastasis associated 1</td>
<td>-0.006</td>
<td>-0.842</td>
</tr>
<tr>
<td>PHE20</td>
<td>PHD finger protein 20</td>
<td>-0.007</td>
<td>-0.577</td>
</tr>
<tr>
<td>XRC1</td>
<td>X-ray repair cross complementing 1</td>
<td>0.000</td>
<td>-0.835</td>
</tr>
<tr>
<td>ABCB1</td>
<td>ATP binding cassette subfamily B member 1</td>
<td>-0.001</td>
<td>-75.564</td>
</tr>
<tr>
<td>TTC27</td>
<td>tetratricopeptide repeat domain 27</td>
<td>-0.001</td>
<td>-4.740</td>
</tr>
<tr>
<td>MYOC</td>
<td>myosin VC</td>
<td>-0.003</td>
<td>12.686</td>
</tr>
<tr>
<td>PRSS11</td>
<td>protease serine 11/HtrA serine peptidase 1 (HTRA1)</td>
<td>0.012</td>
<td>-0.455</td>
</tr>
<tr>
<td>VEGF</td>
<td>vascular endothelial growth factor A</td>
<td>0.009</td>
<td>4.823</td>
</tr>
<tr>
<td>JOSD2</td>
<td>Josephin domain containing 2</td>
<td>0.008</td>
<td>-30.042</td>
</tr>
<tr>
<td>HSPG2</td>
<td>heparan sulfate-glucosamine 3-sulfotransferase 4</td>
<td>0.015</td>
<td>-321.906</td>
</tr>
<tr>
<td>ABCB1</td>
<td>ATP binding cassette subfamily B member 1</td>
<td>-0.001</td>
<td>-75.564</td>
</tr>
<tr>
<td>SLC25A39</td>
<td>solute carrier family 25 member 39</td>
<td>-0.011</td>
<td>-3.054</td>
</tr>
<tr>
<td>DDX58</td>
<td>DExD/H-box helicase 58</td>
<td>-0.010</td>
<td>-1.747</td>
</tr>
<tr>
<td>CISD1</td>
<td>CDGSH iron sulfur domain 1</td>
<td>-0.010</td>
<td>-0.311</td>
</tr>
<tr>
<td>SLC25A39</td>
<td>solute carrier family 25 member 39</td>
<td>-0.011</td>
<td>-3.054</td>
</tr>
<tr>
<td>TNFSF13B</td>
<td>TNF superfamily member 13b</td>
<td>-0.014</td>
<td>-1.663</td>
</tr>
<tr>
<td>PPIA1</td>
<td>PTPRF interacting protein alpha 1</td>
<td>-0.015</td>
<td>0.748</td>
</tr>
<tr>
<td>SCARB2</td>
<td>scavenger receptor class B member 2</td>
<td>-0.015</td>
<td>-0.046</td>
</tr>
</tbody>
</table>

*ddTCip cohort, ovarian cancer patients who were enrolled in this theranostic study and received intraperitoneal carboplatin plus intravenous dose-dense paclitaxel; **TCGA cohort, published data sets in The Cancer Genome Atlas (ovarian cancer patients who received taxane and/or platinum therapy); ***Adj_R2, adjusted R (regression coefficient)-squared; **** p, p-value for the slope; #N/A, not analyzed
To clarify their clinical impact, we conducted Kaplan-Meier analyses in all 76 enrolled ddTCip cases including censored ones and 189 TCGA cases. The analyses demonstrated that the positive expression of SPINK1 alone signified worse PFS in both cohorts when tumors with greater than the mean value of the expression were defined as high expressors (Fig. 2). Although the number of positive expression cases in ddTCip cohort was very few (n = 2), the larger scale validation study in TCGA cohort confirmed the suggested result. SPINK1, even alone, might be an active predictive biomarker of response to ddTCip therapy in patients with suboptimal residual ovarian cancer.

### Development Of Multiplex Prediction Model For Pfs

Nevertheless, drug resistance is multifactorial and multiple genes are involved in the mechanisms. We previously proposed 3 multiplex prediction models using expression data of selected key marker genes (Formulae A, B, and C) with showing high fitness [5, 21]. The expected high predictivities, however, were not confirmed in this study: Formula C alone showed apparent high fitness, but the correlation slope observed in 76 ddTCip patients was opposite to that in 189 TCGA cases (Supplementary Fig. S1).

To develop a truly active prediction model, we sought out highly correlative genes with PFS genome-wide using DNA microarray expression data as the first step. The correlation study in ddTCip cohort yielded a total of 61 correlative genes (p < 0.05; Supplementary Table S3). They were mostly classified into cancer-related genes through the functional enrichment analyses (Supplementary Table S4). They included ABCC2 and TNNT3 but did not contain SPINK1 (p = 0.094).

Multiple regression analysis on the 61 genes provided a novel prediction model (Formula-1) with showing the highest fitness ($R^2 = 0.76, p = 7.5e-11$), which was composed of expression data of 19 effective genes whose expression levels would most correctly explain the value of therapeutic efficacy, PFS (Fig. 3). The predictivity was confirmed also when using TCGA datasets ($R^2 = 0.06, p = 4.0e-4$). To improve the predictivity, we integrated the proteome data in the selection of more active genes. Of the selected 61 genes, protein expression data of 35 genes were available in TCGA cohort. Among 22 genes demonstrated the significant correlations of their expression levels with those in the protein ($p < 0.05$, Pearson test) (Supplementary Table S5). Their expression data were then subjected to multiple regression analysis, and forward stepwise method finally selected 5 genes as the most effective variables from the 22 genes - CUL1 (cullin 1 gene), SLC5A1 (solute carrier family 5 member 1 gene), GPDA1 (glycerol-3-phosphate dehydrogenase 1 gene), PDE3A (phosphodiesterase 3A gene) and VANGL1 (VANGL planar cell polarity protein 1 gene). This led us to another potent prediction formula, Formula-2 ($R^2 = 0.54, p = 46.5e-64$), which was found to show the higher predictive accuracy of PFS ($R^2 = 0.10, p = 46.22e-9$) than that of Formula-1 in TCGA cohort.

In the case of using expression data of the 22 genes, we also adopted random forest method with leave-one-out cross validation to evaluate the variable importance of each selected gene and the potential for the prediction of individual clinical response to platinum and/or paclitaxel chemotherapy. These approaches demonstrated that KRIT1 (Krev interaction trapped protein 1, ankyrin repeat containing gene) expression was the most important variable [Supplementary Fig. S2(A)] in the constructed prediction model for PFS. Even so, the expected predictivity of the model was not superior to those provided in the multiple regression analysis [Supplementary Fig. S2(B), (C)].

### Validation Of Previously Suggested Toxicity Marker

In ddTCip cohort, 75 of 76 enrolled patients (98.7%) experienced grade 3/4 adverse events [20]. Hematological toxicities were the most common adverse events, and problematic grade 3 peripheral neuropathy was observed in 8 patients (10.5%). We investigated 31 polymorphisms of 15 genes known as a potent toxicity marker for taxane and/or platinum therapy (Supplementary Table S2). Genotype-toxicity association analysis (Fisher’s exact test) revealed that variation of ABCB1 rs1045642_A > G was significantly associated with grade 3/4 neutro- ($p < 0.0495$) and thrombo-cytopenia ($p < 0.0001$), while ERCC1 rs11615_A > G variant ($p = 0.0090$) was closely accompanied with peripheral sensory neuropathy in ddTCip cases [Supplementary Fig. S3(A)]. GTEx (V8), a public resource of The Genotype-Tissue Expression project, indicated that these 2 SNPs might cause a decrease of each gene expression in several tissue such as testis and tibial nerve tissue (https://gtexportal.org/home/) [Supplementary Fig. S3(B), (C)]. ABCB1 rs1045642_A > G and ERCC1 rs11615_A > G could be a potent single indicator for toxicity induced by ddTCip therapy in advanced ovarian cancer patients.

### Multiplex Prediction Model For Adverse Events
As with the efficacy, we tackled the development of multiplex prediction model for the adverse events. Classification and Regression Tree (CART) method was first applied by using positive discrimination rate (PDR) and AUC as the estimators of the predictivity. The analysis indicated that CART model using a total of 4 combinations of genotypes could predict the occurrence of 4 serious adverse events induced by ddTCip therapy. The genotypes used in the CART prediction model were, 1) \(ABCB1\) (rs20232582_A > C/T and rs3213619_A > G) and \(XRCC1\) rs25487_T > C for grades 3/4 white blood cell decreased (PDR, 69.803; AUC, 0.702), 2) \(ABCB1\) rs1045642_A > G for grades 3/4 platelet count decreased (PDR, 83.562; AUC, 0.749), 3) \(GSTT1\) null and \(ABCB1\) rs2032582_A > C/T for grades 3/4 anemia (PDR, 73.973; AUC, 0.711), and 4) \(XRCC1\) rs25487_T > C, \(UGT1A1\) rs4148323_G > A and \(ABCB1\) rs2032582_A > C/T for grade 2–4 peripheral sensory neuropathy (PDR, 78.082; AUC, 0.723) (Fig. 4).

Multiple logistic regression analysis also provided novel potent multiplex prediction models composed of the combinations of a total of 15 SNPs on 12 genes for 8 serious adverse events (Table 2). These models demonstrated high predictivity for grade 3/4 neutrophile count decreased (PDF, 89.041; AUC, 0.803), grade 3/4 lymphocyte count decreased (PDF, 87.671; AUC, 0.762), grade 2–4 digestive symptom (PDF, 68.493; AUC, 0.728), and grade 2–4 peripheral motor neuropathy (PDF, 82.192; AUC, 0.797), but the predictivity for grade 3/4 white blood cell decreased was inferior to those in CART models (PDF, 49.315; AUC, 0.586).

### Table 2

<table>
<thead>
<tr>
<th>Gene</th>
<th>Polymorphism ID</th>
<th>Genotype</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>White blood cell decreased (grade 3–4)</td>
<td>Neutrophile count decreased (grade 3–4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCB1</td>
<td>rs1045642</td>
<td>A/G</td>
<td>0.006***</td>
</tr>
<tr>
<td>ABCB1</td>
<td>rs1045642</td>
<td>G/G</td>
<td>0.006***</td>
</tr>
<tr>
<td>ABCB1</td>
<td>rs1128503</td>
<td>A/G</td>
<td></td>
</tr>
<tr>
<td>ABCB1</td>
<td>rs2032582</td>
<td>A/T</td>
<td></td>
</tr>
<tr>
<td>ABCB1</td>
<td>rs2032582</td>
<td>A/C</td>
<td></td>
</tr>
<tr>
<td>ABCB1</td>
<td>rs2032582</td>
<td>C/T</td>
<td></td>
</tr>
<tr>
<td>ABCB1</td>
<td>rs2032582</td>
<td>C/T</td>
<td></td>
</tr>
<tr>
<td>ABCB1</td>
<td>rs4148356</td>
<td>A/A</td>
<td></td>
</tr>
<tr>
<td>ABCB2</td>
<td>rs3740066</td>
<td>C/T</td>
<td></td>
</tr>
<tr>
<td>ABCC2</td>
<td>rs3740066</td>
<td>T/T</td>
<td></td>
</tr>
<tr>
<td>CYP1B1</td>
<td>rs1056836</td>
<td>G/C</td>
<td></td>
</tr>
<tr>
<td>CYP3A5</td>
<td>rs776746</td>
<td>T/C</td>
<td></td>
</tr>
<tr>
<td>CYP3A5</td>
<td>rs776746</td>
<td>C/C</td>
<td>0.134</td>
</tr>
<tr>
<td>ERCC1</td>
<td>rs3212986</td>
<td>C/A</td>
<td></td>
</tr>
<tr>
<td>ERCC1</td>
<td>rs3212986</td>
<td>A/A</td>
<td></td>
</tr>
<tr>
<td>ERCC2</td>
<td>rs313181</td>
<td>T/G</td>
<td></td>
</tr>
<tr>
<td>GSTP1</td>
<td>rs1695</td>
<td>A/G</td>
<td>0.114*</td>
</tr>
<tr>
<td>GSTP1</td>
<td>rs1695</td>
<td>G/G</td>
<td></td>
</tr>
<tr>
<td>GSTT1</td>
<td>Null</td>
<td>Null</td>
<td></td>
</tr>
<tr>
<td>UGT1A1</td>
<td>rs4148323</td>
<td>G/A</td>
<td></td>
</tr>
<tr>
<td>XRCC1</td>
<td>rs1799782</td>
<td>G/A</td>
<td></td>
</tr>
<tr>
<td>XRCC1</td>
<td>rs25487</td>
<td>C/C</td>
<td></td>
</tr>
<tr>
<td>XRCC3</td>
<td>rs861539</td>
<td>G/A</td>
<td></td>
</tr>
</tbody>
</table>

| Sum of response | 48 | 63 | 13 | 17 | 42 | 43 | 14 | 21 |
| Positive discrimination rate (%) | 49.315 | 89.041 | 87.671 | 89.041 | 72.603 | 68.493 | 82.192 | 78.082 |
| Area under the curve | 0.586 | 0.803 | 0.762 | 0.914 | 0.784 | 0.728 | 0.797 | 0.829 |

***, p < 0.01; **, 0.01 ≤ p < 0.05; *, 0.05 ≤ p < 0.10
Discussion

Despite the development of new anticancer agents, the unmet medical needs of advanced or metastatic ovarian cancer patients remain high. We recently indicated that ddTCip therapy is a promising front-line chemotherapy in these patients [20], and here demonstrated novel potent biomarkers and powerful multiplex statistical models for prediction of the clinical response to ddTCip therapy. These may contribute to realize a more effective therapeutic strategy of the therapy, precision medicine.

A variety of candidate genes, gene enrichment signatures and genetic variations have been proposed for the prediction of the efficacy and toxicity of platinum, taxane, and the combination regimens [4–19]. For platinum compounds, they include copper transporting factor encoding genes [17, 18], drug metabolizing enzyme genes [4, 16], and nucleotide excision repair pathway genes [9]. For taxane, they include genes encoding cellular efflux mediating protein including P-glycoprotein MDR1 (ABCB1), ABCC1 and ABCC2, metabolizing enzymes -CYP2C8, CYP3A4 and CYP3A5, and the drug action targets such as tubulin beta (TUBB3) [6, 7]. All these, however, do not always accurately predict individual clinical response to the drugs.

This study proposed 3 potent single prediction markers of ddTCip therapy: They are high SPINK1 expression for the efficacy (PFS), ABCB1 rs1045642_A > G for the grade 3/4 bone marrow toxicity (neutro- and thrombo-cytopenia), and ERCC1 rs11615_A > G for the occurrence of peripheral sensory neuropathy. This may be the first report to demonstrate the high potential of SPINK1 as the single efficacy predictor. Although the relevance of ABCB1 rs1045642_A > G and ERCC1 rs11615_A > G to the toxicity of platinum, taxane, and the combination have been reported in a variety of study [19], the indicated role as a predictor of peripheral neurotoxicity is considered of value in ovarian cancer treatment. The serious neuropathy is often irreversible and distresses the patient extremely.

SPINK1 encodes pancreatic secretory trypsin inhibitor, which is secreted from pancreatic acinar cells into pancreatic juice, and reportedly drives ovarian cancer cell proliferation through activation of EGFR or IL-6 signaling [22, 23]. To identify a truly powerful efficacy predictor, we sought out correlative genes with PFS in the expression level through the analyses in 2 different patient population (ddTCip and TCGA cohorts), targeting 2 different gene groups (41 reported exploratory markers and whole genes), through 2 different correlation analyses with PFS (linear regression analysis in non-censored cases and Kaplan-Meier methods in all patients including censored cases for ddTCip cohort), and finally revealed that the positive expression of SPINK1 alone precisely signified worse PFS in the quantified expression levels, despite the lack of observed correlation in the DNA microarray expression levels. Our robust selection process can affirm the significant role of SPINK1 as a potent exploratory predictive marker. Although the biological roles in cancers remain unclear [24], recent reports have suggested that SPINK1 promote the recurrence risk in prostate and breast cancers [25]. The suppression of a drug resistant determinant, SPINK1, may contribute not only to enhance the therapeutic activity of ddTCip regimen but also to prevent the disease recurrence. Current attention has been focused also on the role of SPINK1 as an attractive novel therapeutic target of cancer [26]. Although TNNT3 and ABCCC2 were also selected as a correlative gene, but Kaplan-Meier analysis did not validate the suggested association with PFS.

Despite of the rapid progress in omics analysis and bioinformatics, genetic polymorphisms of drug metabolizing enzymes, drug transporters, and DNA repair enzymes are still mainstay in the toxicity prediction [27, 28]. We, therefore, sort out the best marker from these known exploratory genomic markers, differed from the exploration of efficacy marker. As well known, ABCB1 relates to drug efflux, while ERCC1 is involved in the repair of DNA lesions formed by platinum compounds. We further suggested here that both single nucleotide polymorphism (SNP) might reduce the corresponding gene expression in normal tissues. This might cause a prolonged drug retention in bone marrow and an inhibition of DNA repair in nerve system, and involve in the occurrence of the adverse events. In fact, not a few reports have suggested such close associations of ABCB1 rs1045642 with paclitaxel-mediated toxicity [29] and of ERCC1 rs11615 genotype with neuropathy in patients with platinum-based chemotherapy [30]. ABCB1 rs1045642_A > G and ERCC1 rs11615_A > G are probably to be a most potent single prediction marker for the toxicity induced by ddTCip therapy.

Even so, particularly important in the current study is the successful construction of multiplex prediction models. It is obvious that drug response mechanism is intricate and multifactorial. A comprehensive molecular view is necessary for understanding the complex mechanisms. We, therefore, applied the multiplex statistics model and machine learning methods to composed powerful multiplex prediction models with showing high fitness.

For the efficacy prediction, our efforts finally yielded a putative multiplex model using expression data of 5 genes-CUL1, SLC5A1, GPD1, PDE3A and VANGL1-. A total of 61 genes were first selected though DNA microarray screening analysis of correlative genes with PFS, then picked more reliable 22 genes though the integrated analysis on mRNA and protein expression using TCGA data sets, and finally narrowed down to 5 genes useful as the most active variables. The 22 genes were mostly involved in the pathways of "Carbohydrate Metabolism" and "Lipid metabolism", in accordance with the reported finding that dynamic proteins transcriptionally regulated in response to nutrient demand or other perturbations are more highly correlated with mRNA in the expression levels in ovarian cancer specimens [31]. Interestingly, all 5 genes finally selected have been proved to play a significant role in cancers. CUL1 and VANGL1 are known as an oncogene [32–34], while GPD1 acts as a tumor suppressor [35, 36], PDE3A and SLC5A1 are well recognized as a valuable therapeutic target of cancers [37–41]. These suggest that genome-wide screening and consecutive integrated proteo-transcriptomic selection would work effectively in the determination of key variables in the prediction model. The observed difference in the selected genes from single predictive markers might reflect the intricate multiplex network of drug response mechanisms, as we intended. Even so, our previous prediction models constructed in the analysis of the limited number of cases was shown not to be active at all. This apparently indicates that the practical usefulness needs to be evaluated by a larger prospective study.

As for the toxicity, CART analyses demonstrated that 4 combinations of genetic variants on 4 genes could predict the occurrence of 4 serious adverse events: Grades 3/4 white blood cell decreased, Grades 3/4 platelet count decreased, Grades 3/4 anemia, Grade 2–4 peripheral sensory neuropathy. Multiple logistic regression analyses provided another potent multiplex prediction models composed of the combinations of a total of 15 genetic
variants on 12 genes (ABCB1 rs1045642, rs1128503 and rs2032582; ABCC1 rs4148356; ABCC2 rs3740066; CYP1B1 rs1056836; CYP3A5 rs776746; ERCC1 rs3212986, ERCC2 rs13181; GSTP1 rs1695, GSTT1 null, UGT1A1 rs4148323, XRCC1 rs1799782 and rs25487; and XRCC3 rs861539) for 8 serious adverse events including not only bone marrow suppression but also digestive symptoms and peripheral motor neuropathy. All prediction models demonstrated high positive discrimination rates (PDR) and area under the curve (AUC) values sufficient in the toxicity prediction.

As already described, all variables used in the models have been proved to be functionally involved in the drug response mechanisms. Even so, the limited understanding of the biology of proposed markers and the interaction of the selected variables in the prediction models will become a hurdle to the development of novel therapeutics. A set of several gene expressions or genetic variations likely works complementary in our prediction models, which suggest that some interactions of each variable may exist in the molecular level but even for the outline remains unknown. Although our prediction models demonstrated the advantage in prediction of PFS, prediction of overall survival is also our interests. For the toxicity prediction, the ethnic difference also should be integrated into the utility assessment. ERCC2 rs13181 and CYP3A4*1B SNPs have been often reported to be associated with paclitaxel-induced cytotoxicity [42, 43], but the observed minor allele frequencies of ERCC2 rs13181 and CYP3A4*1 genotype in Japanese significantly differed from those of Caucasian and African population published in 1000 Genomes database. In fact, in the ddTCip cases, the minor allele frequency of ERCC2 rs13181 was only 0.047, quarter to one-eight frequencies compared to the European, American, or African population.

The comparative advantage of intraperitoneal carboplatin plus paclitaxel (ddTCip therapy) to intravenous chemotherapy remains undetermined [44], although recent Phase III study on ddTCip therapy demonstrated the improved progression-free survival in patients with ovarian, fallopian tube, or primary peritoneal carcinoma [45]. Despite the most active modality is still controversy, the fact remains that the paclitaxel plus carboplatin (TC) regimen is an essential treatment in patients with intractable advanced ovarian cancers. Our proposed prediction models might contribute to the therapeutic outcome through the section of suitable patients for the treatment. We are now planning a larger-scale prospective clinical study to confirm the clinical significance of the predictive markers and prediction models along with continuing our search for the functional roles of the selected markers and their interactions in drug sensitivity and toxicity.

Conclusion

This theranostic study on a promising first-line chemotherapy, ddTCip therapy, demonstrated several potent single biomarkers: SPINK1 expression for the efficacy, and ABCB1 and ERCC1 polymorphisms for the toxicity. We further successfully composed robust multiplex prediction models for the individual responses: Efficacy prediction model using expression data of 5 genes, and CART models using a total of 4 genotype combinations and multiple regression models using 15 polymorphisms on 12 genes for the prediction of the adverse events. These prediction models were proved to function well in the validation analyses using TCGA data sets. SPINK1 appeared to be a critical resistant determinant of ddTCip therapy, which may indicate that SPINK1 could be a novel therapeutic target. These may raise the potential to realize a precision medicine in the essential treatment and address an unmet medical need in ovarian cancer treatment.

Abbreviations

ddTcip, intraperitoneal carboplatin in combination with dose-dense paclitaxel; PFS, progression free-survival; CART, Classification and Regression Tree; SPINK1, Serine Peptidase Inhibitor Kazal Type 1; ABCB1, ATP Binding Cassette Subfamily B Member 1; ERCC1, Excision Repair Cross-Complementation Group 1; FIGO, the International Federation of Gynecology and Obstetrics; IP, Intraperitoneal administration; ECOG, Eastern Cooperative Oncology Group; TXL, paclitaxel; CBDCA, carboplatin; RIN, RNA integrity number; RT-PCR, reverse transcription polymerase chain reaction; ACTB, actin beta; ABCB1, ATP Binding Cassette Subfamily C Member 1; ABCC2, ATP Binding Cassette Subfamily C Member 2; CYP1B1, Cytochrome P450 Family 1 Subfamily B Member 1; CYP2C8, Cytochrome P450 Family 2 Subfamily C Member 8; CYP3A4, Cytochrome P450 Family 3 Subfamily A Member 4; CYP3A5, Cytochrome P450 Family 3 Subfamily A Member 5; ERCC2, Excision Repair Cross-Complementation Group 2; GSTP1, Glutathione S-Transferase Pi 1; UGT1A1, UDP Glucuronosyltransferase 1 family, polypeptide A1; XRCC1, X-Ray Repair Cross Complementing 1; XRCC3, X-Ray Repair Cross Complementing 3; SNP Single Nucleotide Polymorphism; GSTM1, Glutathione S-transferase M1; GSTT1, Glutathione S-Transferase Theta 1; TCGA, The Cancer Genome Atlas; AIC, Akaike information criterion; VIF, Variance Inflation Factor; AUC, Area Under the Curve; PDF, Positive Discrimination Rate; IPA, Ingenuity Pathway Analysis; TNNT3, Troponin T3, Fast Skeletal Type; IRF9, Interferon Regulatory Factor 9; CUL1, cullin 1; SLC5A1, solute carrier family 5 member 1; GPD1, glycerol-3-phosphate dehydrogenase 1; PDE3A, phosphodiesterase 3A; VANGL1, VANGL planar cell polarity protein 1; KRIT1, Krev interaction trapped protein 1; GTEx, The Genotype-Tissue Expression project;

Declarations

Acknowledgments

We would like to thank the patients who enrolled in the study. We also thank Dr. Satoru Wada, Ms. Keiko Kurihara and Ms. Azusa Sekine for their helpful technical assistance, and Research Institute for Development Therapeutics and Division of Development of Therapeutics, Research Center for Genomic Medicine, Saitama Medical University for the equipment support.

Authors’ contributions
RK-I, NI, HE, KH, KF, and MN made substantial contributions to conception and design., RK-I, NI, MS, ST, HF, HE, TO, TS, MS, and KH participated in the patient recruitment, sample, and data collection. RK-I, KS, JC, and MN were responsible for analyses and interpretation of data. All authors have been involved in drafting the manuscript or revising it critically for important intellectual content. All authors have given final approval of the version to be published and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**Funding**

This work was financially supported by the Development Organization for Frontier Medical Education and Therapeutics in Japan (DOFMET-08 to KF and MN), by Health, Labor and Welfare Science Research Grant from the Japan Ministry of Health, Labour and Welfare [H21-3rd Comprehensive 10-year Strategy for Cancer Control-010 to MN], by the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan, and Promotion Plan for the Platform of Human Resource Development for Cancer, the Fostering Health Professionals for Changing Needs of Cancer and New Paradigms – Establishing Center for Fostering Medical Researchers of the Future Programs to Saitama Medical University and Gunma University.

**Availability of data and materials**

The authors declare that the data supporting the findings of this study are available within the paper and its supplementary files. Additional data are available from the corresponding author upon reasonable request.

**Ethics approval and consent to participate**

This study was approved by the Institutional Review Board at Gunma University, Saitama Medical University, Tottori University School of Medicine, Iwate Medical University and Jichi Medical University, which was registered at UMIN Clinical Trials Registry (ID: UMIN000001713) on Feb 16th, 2009. All women provided written informed consent to participate including donation of their peripheral blood and surgical specimens before inclusion in the study. The study was conducted in accordance with Japanese guidelines and regulations and the Declaration of Helsinki.

**Consent for publication**

Not applicable

**Competing interests**

The authors have declared that no conflict of interest exists. All authors have read the journal's policy on disclosure of potential conflicts of interest and agreed to the journal's authorship statement.

**References**


Figures

Fig. 1

**Figure 1**

Figure legend not available with this version.
Figure 2

Figure legend not available with this version.
Figure 3

Figure legend not available with this version.
Figure 4

Figure legend not available with this version.

Supplementary Files

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

- SupplementaryFig.S1.pptx
- SupplementaryFig.S2.pptx
- SupplementaryFig.S3.pptx
- SupplementaryTableS1.xlsx
- SupplementaryTableS2.xlsx
- SupplementaryTableS3.xlsx
- SupplementaryTableS4.xlsx
- SupplementaryTableS5.xlsx