Early Modeled Longitudinal CA-125 Kinetics and Survival of Ovarian Cancer Patients: A GINECO AGO MRC CTU Study

Olivier Colomban1, Michel Tod1,2, Alexandra Leary3,4, Isabelle Ray-Coquard5, Alain Lortholary6, Anne Claire Hardy-Bessard7, Jacobus Pfisterer8, Andreas Du Bois9, Christian Kurzeder10,11, Alexander Burges12, Julien Péron3,13, Gilles Freyer1,12, and Benoit You1,3,12

Abstract

Purpose: Regarding cancer antigen 125 (CA-125) longitudinal kinetics during chemotherapy, the actual predictive value of the Gynecologic Cancer Intergroup (GCIG) CA-125 response criterion is questioned. The modeled CA-125 elimination rate constant KELIM exhibited higher prognostic value in patients with recurrent ovarian cancer enrolled in the CALYPSO trial. The objective was to validate the higher predictive and prognostic values of KELIM during first-line treatments.

Experimental Design: Data from three large phase III trials were analyzed: AGO OVAR 9 (learning set: carboplatin-paclitaxel (CP) ± gemcitabine; n = 1,288); AGO OVAR 7 (validation set: CP ± topotecan; n = 192); and ICON7 (validation set: CP ± bevacizumab; n = 1,388). The CA-125 profiles were fit with a nonlinear mixed-effect model during the first 100 days, and the individual KELIM were calculated. KELIM prognostic and predictive values for survival were assessed against GCIG criterion and other prognostic factors in univariate/multivariate analyses.

Results: The GCIG CA-125 endpoint provided no meaningful predictive/prognostic information. C-index analyses confirmed the higher predictive value of KELIM compared with GCIG criterion for progression-free survival and overall survival (OS). KELIM provided reproducible prognostic information. Patients with favorable KELIM ≥ upper tercile (0.0711 per days) consistently experienced better OS, with HRs between 0.44 and 0.58 (e.g., median OS >65 months vs. <35 months).

Conclusions: Modeled KELIM provides higher predictive and prognostic information based on CA-125 longitudinal kinetics compared with GCIG response criteria during first-line chemotherapy. Integration of this endpoint in guidelines may be considered. Individual KELIM and survival simulations can be calculated at http://www.biomarker-kinetics.org/. Further assessment of the surrogate value of KELIM treatment–related variations in a GCIG meta-analysis is warranted.

Introduction

Approximately 75% of newly diagnosed ovarian carcinomas are associated with peritoneal involvement known to be poorly assessable using conventional imaging (1). Thus, the prognostic value of cancer antigen 125 (CA-125) kinetics has been extensively analyzed as a way of overcoming the limitations of imaging (2, 3). In 2004, the Gynecologic Cancer Intergroup (GCIG) defined the CA-125 response as a 50% reduction in CA-125 levels maintained for at least 28 days (4). However, Lee and colleagues recently reported that this CA-125 response criterion was a poor predictor of treatment benefit in patients with recurrent ovarian cancers using data from the CALYPSO phase III trial (5). Indeed the two time-point-based kinetic strategies are limited by high inter- and intra-individual variability, and may thus not be extrapolatable to other patient populations, or treatment settings, than those assessed initially (2).

Strategies based on longitudinal analyses of serum tumor marker dynamics during treatment, using mathematical modeling, have been developed to define the equations describing serum tumor kinetics, thereby reducing the impact of this variability (2, 6–10). Moreover, prognostic modeled kinetic parameters, which are expected to characterize the decline curves, can subsequently be extracted to discriminate favorable and
Translational Relevance

There is a need for easily assessable intermediate endpoints in early-phase trials for guiding the development of novel anticancer agents in ovarian cancers. Cancer antigen 125 (CA-125) longitudinal kinetics during chemotherapy is an indicator of tumor chemosensitivity, but the actual predictive/prognostic values of the Gynecologic Cancer Intergroup (GCIG) CA-125 response criterion are questioned. The same mathematical model could be successfully used for characterizing the CA-125 kinetics of patients enrolled in three large randomized clinical trials testing different first-line chemotherapy regimens. The strength of the statistical association of overall survival with the modeled CA-125 elimination rate constant KELIM was consistently stronger than with GCIG CA-125 criterion, thereby supporting its potential replacement in guidelines. Patient-individual KELIM can easily be calculated online and associated with survival predictions. The role of KELIM treatment–related change as a potential relevant surrogate endpoint of survival benefit will be explored in a GCIG meta-analysis.

unfavorable dynamical profiles. This strategy was promising for different serum tumor marker kinetics, such as CA-125, PSA, hCG, AFP, and circulating tumor cells (2). In a retrospective study of the CALYPSO trial, the modeled CA-125 KELIM parameter, meaning CA-125 elimination rate constant, was a significant independent prognostic factor for progression-free survival (PFS) in patients with platinum-sensitive recurrent ovarian cancers (10). KELIM is a modeled kinetic parameter linked to tumor marker elimination rate during systemic treatment, which is independent of renal function, and can be assimilated to CA-125 clearance. The higher the KELIM, the faster the elimination, the greater the chemotherapy effect, and thus, better prognosis.

The population kinetic approach requires prior validation of the population model parameters, along with the inter- and intra-individual variability, before Bayesian individual KELIM can be estimated on the basis of observed patient CA-125 concentration-time changes. Compared with non-Bayesian methods (i.e., fitting the model to patient data without population model), the Bayesian approach requires lower numbers of concentration measurements for estimating KELIM, with narrower confidence intervals (CI).

Modeled KELIM may be more appropriate for analyzing longitudinal kinetics of CA-125 than GCIG criteria. A European collaboration was launched to assess the predictive and the prognostic values of KELIM, against GCIG CA-125 criterion and other prognostic factors, using the datasets of three large randomized phase III clinical trials (AGO OVAR 7; AGO OVAR 9; and ICON-7), where the standard first-line carboplatin-paclitaxel (CP) doublet regimen was compared with triplet experimental arms, including one trial with positive survival effects (11–13). This work was needed because the model parameters developed for estimating KELIM in patients with recurrent ovarian cancer enrolled in the CALYPSO study had to be readjusted to the CA-125 kinetics of patients treated in first-line setting (10).

Materials and Methods

Inclusion criteria

This retrospective analysis included the data from three large multicenter randomized controlled trials evaluating different first-line regimens in patients with advanced ovarian cancers: AGO OVAR 9 (ref. 11; NCT00052468) comparing carboplatin AUC 5 min/mg/L + paclitaxel 175 mg/m² (CP regimen) with/without gemcitabine 800 mg/m² on days 1 and 8, every 3 weeks for six cycles; AGO OVAR 7 (ref. 12; NCT00102375) comparing the CP regimen for six cycles with/without four subsequent cycles of topotecan 1.25 mg/m² on days 1–5 every 3 weeks; ICON-7 (ref. 13; NCT00483782) comparing the CP regimen with/without bevacizumab 7.5 mg/kg given concurrently on day 1 every 3 weeks for five or six cycles and then continued for 12 additional cycles or until progression of disease. The patient eligibility criteria and study designs have been described in detail elsewhere (11–13). The individual patient databases were obtained from collaborative groups after reviewing this study proposal.

In these trials, the CA-125 concentrations were measured in every 3-week cycle. Their primary objectives were to demonstrate overall survival (OS) gains in the experimental arms, in addition to PFS in the ICON-7 trial. The AGO OVAR 7 trial failed to find any benefit in terms of OS with topotecan administration following CP treatment (median: 43.1 months for the experimental arm vs. 44.5 months for CP alone; HR = 1.01; 95% CI, 0.86–1.18; P = 0.88), nor did the AGO OVAR 9 trial for gemcitabine combined with the CP regimen (median: 49.5 months for the triplet arm vs. 51.5 months; HR = 1.05; 95% CI, 0.91–1.20; P = 0.51; refs. 11, 12). However, a PFS benefit was reported with the addition of bevacizumab to CP in the ICON-7 trial (median: 19.0 months for the experimental arm vs. 17.3 months; HR = 0.81; 95% CI, 0.70–0.94; P = 0.004; ref. 13), along with better OS in the prespecified subgroup of high-risk patients reported by Oza and colleagues (39.7 months vs. 30.2 months; HR = 0.78; 95% CI, 0.63–0.97; P = 0.03; ref. 14).

Mathematical modeling of longitudinal CA-125 kinetics and estimation of KELIM

The mathematical modeling of early CA-125 kinetics has been previously described for the CALYPSO trial dataset analysis (10). Individual CA-125 data were analyzed using a population kinetic approach with a nonlinear mixed-effect model. Basic details about the population kinetic approach, which enables estimation of model parameters under sparse sampling conditions with a few time points per patient, are presented in the Supplementary Methods. To eliminate right-skewness in the data distribution and to normalize the CA-125 titer distribution, data were log-transformed. A semimechanistic model based on a kinetic-pharmacodynamic approach was fitted to serum CA-125 values measured during the first 100 treatment days. This time frame was arbitrarily selected to identify early predictive factors of efficacy, while maximizing the chances to obtain a minimum of three CA-125 values, which was required to assess CA-125 kinetics by the population modeling process.

Treatment kinetics were described by a two-compartment model: central compartment (C1) receiving chemotherapy dosing (doses set to 1) and a transit compartment (C2) to describe the treatment lag-time effect (Supplementary Fig. S1; ref. 15). The CA-125 production inhibition induced by the treatment is
expressed by an indirect effect model using an E_{50} (E_{50}) relationship (16). This model was described by the following equations:

1. \frac{dC_1}{dt} = \frac{K}{C_2} C_1
2. \frac{dC_2}{dt} = \frac{K{PROD}}{C_1} E_{\text{EFFECT} E_{50}} - \frac{K_{\text{ELIM}}}{C_2} C_1
3. \frac{dC_{A125}}{dt} = \frac{K_{\text{PROD}}}{C_2} E_{\text{EFFECT} E_{50}} / C_0 - \frac{K_{\text{ELIM}}}{C_2} C_{A125}
4. E_{\text{EFFECT} E_{50}} = 1 - \frac{C_1}{C_0 + C_2}

The initial conditions at time 0 were:

1. (1) C_1(0) = 0
2. (2) C_2(0) = 0
3. (3) C_{A125}(0) = C_{A0}

Where K is the treatment kinetic rate constant (per days), \text{KPROD} is the CA-125 tumor production rate (IU/mL/days), E_{50} is the concentration producing 50% of the maximum effect (AU), K_{\text{ELIM}} is the CA-125 elimination rate (per days), and CA_{0} is the estimated CA-125 at time = 0 (IU/mL).

The AGO OVAR 9 dataset, which included the greatest number of patients, was used as a learning dataset to adjust the parameters of the model initially built for the CALYPSO trial to first-line treatment trial findings. The AGO OVAR 7 and ICON-7 trial datasets were used as external validation sets: the AGO OVAR 9 final model (final population parameters) was used to calculate individual KELIM values in AGO OVAR 7 and ICON-7 trial patients as empirical Bayesian estimates.

Model qualification

The CA-125 predictions, obtained with the final model built with the AGO OVAR 9 database and implemented for the three datasets, were qualified (17). SEs of estimated parameters and goodness-of-fit plots (i.e., plots of observations vs. predictions and of the distribution of normalized prediction errors) were used as major criteria (18). Moreover, 500 replicates of all individual CA-125 decline profiles were simulated using the final population model to perform a visual predictive check (VPC): the 10th, 50th, and 90th percentiles of the observed CA-125 values were compared with the 95% CI computed from the 500 simulated replicates (19). The observed and simulated median versus time profiles were visually compared.

Assessment of KELIM predictive value with respect to other known prognostic factors

The discriminatory predictive ability of KELIM regarding PFS and OS were assessed using the C-index in univariate and multivariate analyses of every trial dataset separately and in all trials pooled together (20, 21). The following previously reported prognostic factors were also tested (14, 22): the CA-125 response according to the GCIG definition (Rustin criteria, decline of CA-125 by 50% over a minimum 28 day period: favorable vs. unfavorable; ref. 4); disease stage (I–II, III, or IV); and histology (clear cell carcinoma vs. other) for the AGO OVAR 9 and 7 datasets. Moreover, the tumor grade (1 and 2 vs. 3) and the high-risk status defined by Oza and colleagues [ref. 14; including stage IV disease, inoperable stage III disease, or suboptimally...
debunked (>1 cm) stage III disease, along with treatments arms, were tested in the ICON-7 dataset. Empirically based, the 95% CI of C-index was estimated using a bootstrapping method (Q2.5% and Q97.5% of 1,000 resampling).

Furthermore, the “trial level association” between KELIM and treatment efficacy was explored by assessing a nonweighted linear regression between KELIM median value ratios (e.g., ratio of median KELIM value in arm A/median KELIM value in arm B) with the corresponding PFS and OS Cox proportional HRs (e.g., HR for OS in arm A compared with arm B) in the three trials (23).

Assessment of the KELIM prognostic value with respect to other known prognostic factors using nonparametric survival tests

To assess the magnitude of the relationships between KELIM and survival, the prognostic value of KELIM categorized by the terciles found in the learning set AGO OVAR 9 was tested using univariate and multivariate analyses for PFS and OS in this dataset, and then in the two external validation datasets. Univariate survival analyses were performed using the Kaplan–Meier method and log-rank tests. All covariates found positively significant in univariate analyses were integrated in a multivariate analysis using a Cox proportional hazards regression model, and the HR and 95% CI were computed. The final survival model was obtained using a backward selection.

A parametric model linking KELIM and OS for predicting patient survival based on estimated individual KELIM

An accelerated failure time (AFT) parametric model was built to relate KELIM and OS. The selection of best model among those tested, including Weibull, normal, exponential, log-logistic, etc., was based on the Akaike information criterion. KELIM was first considered as a categorized covariate using the same terciles as found in AGO OVAR 9 study to see whether the outcomes of the parametric and the nonparametric models were consistent.

Second, KELIM was considered as a continuous covariate standardized by the median found in the three datasets combined together. The objective was to develop a model that could be implemented on the website http://www.biomarker-kinetics.org/ for predicting patient survivals based on calculated individual KELIM, using the AFT model. AFT was used to assess the survival outcomes and their CIs with 1,000 replications of the simulation.

All nonparametric and parametric survival analyses were implemented with a landmark timepoint set at 100 days, because CA-125 was modeled from day 0 to 100. Indeed exclusion of the early progressions observed during the first 100 days avoided the biases related to the links between early progressions and early CA-125 kinetics (20).

Computing process

All tests were implemented using software, including Weibull, normal, exponential, log-logistic, etc., based on the Akaike information criterion. The XPOSE4 program was used for graphical evaluation of model fits (25). Survival analyses, concordance probability (C-index), and AFT analyses were, respectively, obtained using coxph, cindex function (package survreg, WeibullReg (package SurvRegCensCov), and flexsurv (package flexsurv; ref. 26) in R software version 3.4.4.

Results

Patient selection

The data from 2,868 patients fulfilling the CA-125–based criteria were analyzed for modeling purposes: 1,288 patients from the AGO OVAR 9 trial (of 1,741 enrolled patients), 192 patients from the AGO OVAR 7 trial (of 1,308 enrolled patients), and 1,388 from the ICON-7 trial (of 1,528 enrolled patients; Fig. 1). The data from 2,832 patients could be assessed for PFS, and those of 2,860 patients for OS (Fig. 1). The characteristics of these patients are presented in Supplementary Table S1.

Model qualification

Typical parameter estimates, along with the qualification analyses from the final semimechanistic model, are presented in the Supplementary Methods (Supplementary Table S2, Supplementary Figs. S2 and S3).

The interindividual coefficient of variation of KELIM was large (56%). Relative standard errors of KELIM typical values and interindividual coefficient of variation, which represent the estimation precision, were low (2.5% and 2.6%, respectively), as was the shrinkage, thereby suggesting limited risks of biased individual estimates of KELIM. The goodness-of-fit plots showed that individual CA-125 profiles were properly fit during the first 100 treatment days by the model (Supplementary Fig. S2). Although the CIs of the percentiles are narrow due to the high number of patients included in the analysis, the VPC did not reveal invalidation, suggesting correct predictive performance of the model (Supplementary Fig. S3).

When applied to both external validation datasets, goodness-of-fit plots figures, as well as VPC, showed similar outcomes (Supplementary Fig. S3).

Predictive value of KELIM with respect to other prognostic factors, including the GCIG response criteria

The results regarding the respective C-index values of KELIM and other available prognostic factors for PFS and OS are presented in Table 1 and Supplementary Fig. S4. The predictive value of KELIM was consistently higher than GCIG response criteria (e.g., discriminatory ability for OS, 0.60, 95% CI, 0.58–0.62 for KELIM; 0.49, 95% CI, 0.48–0.52 for GCIG criteria). Moreover, KELIM addition in the multivariate C-index model increased the relative apparent performance by 9.7% for PFS, and 8.2% for OS, while addition of GCIG response criteria offered no benefit to the predictive multivariate model. Similar outcomes were noted in every trial considered separately, acknowledging the

<table>
<thead>
<tr>
<th>Covariates</th>
<th>C-index (95% CI)</th>
<th>C-index (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univariate models</td>
<td>KELIM 0.61 (0.59–0.62)</td>
<td>0.60 (0.58–0.62)</td>
</tr>
<tr>
<td>FIGO stage</td>
<td>0.62 (0.61–0.63)</td>
<td>0.61 (0.59–0.62)</td>
</tr>
<tr>
<td>GCIG CA-125 response</td>
<td>0.51 (0.50–0.52)</td>
<td>0.49 (0.48–0.52)</td>
</tr>
<tr>
<td>Multivariate models</td>
<td>KELIM + FIGO stage 0.68 (0.66–0.69)</td>
<td>0.66 (0.64–0.68)</td>
</tr>
<tr>
<td>KELIM + GCIG CA-125 response 0.62 (0.60–0.63)</td>
<td>0.60 (0.58–0.62)</td>
<td></td>
</tr>
<tr>
<td>FIGO + GCIG CA-125 response 0.62 (0.60–0.64)</td>
<td>0.60 (0.60–0.63)</td>
<td></td>
</tr>
<tr>
<td>KELIM + FIGO stage + GCIG CA-125 response 0.68 (0.66–0.69)</td>
<td>0.66 (0.64–0.68)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: FIGO, Federation Internationale des Gynaecologistes et Obstetrites.
lower statistical powers of these analyses due to the smaller numbers of patients (Supplementary Table S3).

Linear regression explorations between treatment-related KELIM ratios and survival HRs suggested the potential surrogate value of KELIM ratios: slope, 1.20, 95% CI, −0.44–2.85; intercept, −0.24, 95% CI, −2.10–1.62 with \( R^2 = 0.99 \) for OS (Supplementary Fig. S5).

### Prognostic value of KELIM with respect to other prognostic factors, including the GCIG response criteria, using nonparametric tests

To illustrate the association strength between survival and KELIM, the terciles of KELIM found in AGO OVAR 9 dataset (0.0482 per days and 0.0711 per days/day) were tested as the prognostic cut-off values. Strong prognostic values for PFS and OS were found in this dataset (Table 2; Supplementary Table S4). For example, median OS were not reached (＞75.0), 52.0, and 32.6 months in patients with favorable, intermediate, and unfavorable KELIM, respectively (\( P < 0.001 \); Fig. 2). Multivariate Cox analyses confirmed the independent and strong prognostic value of KELIM for PFS and OS when tested with other assessed prognostic factors, including GCIG CA-125 response criterion (e.g., for OS, HR = 0.62, 95% CI, 0.51–0.74 for intermediate KELIM tercile; and HR = 0.45, 95% CI, 0.37–0.55 for favorable KELIM tercile, compared with unfavorable KELIM tercile, \( P < 0.001 \); Table 2; Supplementary Table S4).

Similar outcomes were found in both external validation datasets with the same KELIM thresholds: AGO OVAR 7, median OS = 64.2, 46.7, and 34.3 months for favorable, intermediate, and unfavorable KELIM, respectively, \( P = 0.01 \); ICON-7: median OS = 65.0, 61.5, and 33.4 months for favorable, intermediate, and unfavorable KELIM, respectively, \( P < 0.001 \). When assessed with other significant prognostic factors in multivariate analyses, KELIM was a significant independent and consistent prognostic factor of PFS and OS in both datasets (Table 2; Supplementary Table S4).

The prognostic value of KELIM for OS, adjusted for the treatment arms, was assessed in the three datasets (Table 3; Supplementary Fig. S6). In all trials, the best median survivals were observed in patients with favorable KELIM regardless of the treatment arm, including in the positive ICON-7 trial. Further exploratory analyses were performed using the ICON-7 trial data, with adjustment for risk classes defined by Oza and colleagues (14). As shown in Supplementary Table S5, the strong prognostic value of KELIM was maintained within these prognostic risk classes. Patients with a favorable KELIM consistently experienced longer PFS and OS than patients with unfavorable KELIM, whether they were randomized to bevacizumab or placebo. Bevacizumab was consistently associated with better median survival, but it did not compensate for the impact of unfavorable KELIM on patient survival.

### Prognostic value of GCIG CA-125 response criteria

The GCIG CA-125 response criterion was significantly, but inversely associated with PFS using multivariate analysis in the ICON-7 trial only (\( HR = 1.28 \) if GCIG favorable response; 95% CI, 1.09–1.50; Supplementary Table S4). It was not significantly associated with OS in multivariate analyses (Table 2).

#### The parametric model linking KELIM and OS

The best parametric model was based on a Weibull distribution, where KELIM was considered as a categorized or a continuous covariate. The estimated parameters and outcomes of these models are presented in Supplementary Table S6. The gradual prognostic value of KELIM terciles found with the nonparametric model was confirmed with the parametric model (Fig. 3A).

The AFT model with continuous KELIM was used to predict patient survival based on calculated KELIM (standardized by the

---

**Table 2. Covariates associated with OS (months) in the three datasets**

<table>
<thead>
<tr>
<th>Covariates</th>
<th>Classes</th>
<th>Univariate OS</th>
<th>Multivariate OS</th>
<th>Univariate OS</th>
<th>Multivariate OS</th>
<th>Univariate OS</th>
<th>Multivariate OS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AGO OVAR 9 (N = 1,289)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Univariate P</td>
<td>HR (95% CI)</td>
<td>Univariate P</td>
<td>HR (95% CI)</td>
<td>Univariate P</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td>KELIM terciles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unfavorable</td>
<td>[min–Q33%]</td>
<td>32.6</td>
<td>&lt;0.001 REF</td>
<td>34.3</td>
<td>0.05 REF</td>
<td>33.4</td>
<td>&lt;0.001 REF</td>
</tr>
<tr>
<td>Intermediate</td>
<td>[Q33%–Q66%]</td>
<td>52.0</td>
<td>0.62 (0.51–0.74)</td>
<td>46.7</td>
<td>0.53 (0.31–0.92)</td>
<td>61.5</td>
<td>0.51 (0.42–0.61)</td>
</tr>
<tr>
<td>Favorable</td>
<td>[Q66%–max]</td>
<td>NR</td>
<td>0.45 (0.37–0.55)</td>
<td>64.2</td>
<td>0.58 (0.35–0.97)</td>
<td>65.0</td>
<td>0.44 (0.35–0.53)</td>
</tr>
<tr>
<td>Arm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard</td>
<td>50.0</td>
<td>0.83</td>
<td>64.2</td>
<td>0.21</td>
<td>54.7</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>48.0</td>
<td>0.83</td>
<td>41.9</td>
<td>0.83</td>
<td>53.9</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>FIGO stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I–II</td>
<td>NR</td>
<td>&lt;0.001 REF</td>
<td>NR</td>
<td>&lt;0.001 REF</td>
<td>NR</td>
<td>&lt;0.001 REF</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>46.2</td>
<td>7.03 (3.32–6.62)</td>
<td>47.6</td>
<td>2.79 (0.84–9.23)</td>
<td>48.2</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>26.6</td>
<td>7.82 (5.38–11.35)</td>
<td>29.4</td>
<td>6.55 (1.88–22.80)</td>
<td>32.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clear cell</td>
<td>NR</td>
<td>0.30</td>
<td>NR</td>
<td>0.14</td>
<td>61.5</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>49.2</td>
<td>0.30</td>
<td>42.3</td>
<td>0.30</td>
<td>53.1</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>GCIG CA-125 response</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unfavorable</td>
<td>49.4</td>
<td>0.50</td>
<td>46.7</td>
<td>0.24</td>
<td>61.5</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Favorable</td>
<td>49.6</td>
<td>0.50</td>
<td>42.5</td>
<td>0.50</td>
<td>50.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 + 2</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td>53.9</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td>54.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td>65.0</td>
<td>&lt;0.001 REF</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>NA</td>
<td></td>
<td>NA</td>
<td></td>
<td>32.4</td>
<td>2.22</td>
<td></td>
</tr>
</tbody>
</table>

**NOTE:** KELIM AGO OVAR 9 Q33% = 0.0482 per days; Q66% = 0.0711 per days.

Abbreviations: FIGO, Federation Internationale des Gynaecologistes et Obstetristes; NA, not available in the datasets; NR, not reached; REF, reference.
median value of KELIM 0.06 per days in the three pooled data-sets). As an illustration, Fig. 3B displays survival predictions for different values of standardized KELIM that were arbitrarily set up (KELIM/0.06 = 0.5; = 1; or = 1.5), with respect to the observed Kaplan–Meier curve. There was high consistency in between the observed survival curve with the simulated curve at KELIM/0.06 = 1, thereby suggesting the concordance of simulated with observed survivals.

**Discussion**

The therapeutic escalation strategies currently implemented in first-line setting of patients with ovarian cancer, and offered by the emergence of many novel anticancer agents (antiangiogenic drugs, immunotherapies, DNA repair interacting agents etc.), are limited by the lack of effective predictive companion biomarkers (27). There is a high need for intermediate predictive endpoints, easily assessable during early-phase trials, and able to inform on the benefits to expect with novel drug candidates or strategies, before launching large and expensive phase III trials (20, 27).

CA-125 longitudinal decline during treatment may provide information about chemo-sensitivity and treatment-related benefits (27). Although the Rustin and colleagues CA-125 decline criterion has been officially recognized as the GCIG response criteria for drug development, its actual predictive has recently been questioned (4, 5), thereby warranting development of more modern approaches with current bioinformatics facilities.
Table 3. Impact of KELIM categorized by terciles (favorable, intermediate, and unfavorable) on PFS and OS adjusted for treatments

<table>
<thead>
<tr>
<th>Treatment arm</th>
<th>Favorable* KELIM</th>
<th>Intermediate* KELIM</th>
<th>Unfavorable* KELIM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median PFS (months; 95%CI)</td>
<td>Median PFS (months; 95%CI)</td>
<td>Median PFS (months; 95%CI)</td>
</tr>
<tr>
<td>AGO OVAR 9 CP</td>
<td>29.8 (23.0–42.9)</td>
<td>18.7 (15.3–24.3)</td>
<td>10.4 (9.1–11.6)</td>
</tr>
<tr>
<td>AGO OVAR 9 CP + gemcitabine</td>
<td>29.7 (22.0–36.2)</td>
<td>16.5 (13.8–20.1)</td>
<td>9.7 (8.7–10.8)</td>
</tr>
<tr>
<td>AGO OVAR 7 CP</td>
<td>27.7 (11.9–NR)</td>
<td>19.7 (14.6–21.9)</td>
<td>12.0 (8.8–31.5)</td>
</tr>
<tr>
<td>AGO OVAR 7 CP followed by topotecan</td>
<td>16.7 (9.8–NR)</td>
<td>14.6 (11.9–28.9)</td>
<td>9.1 (7.3–12.3)</td>
</tr>
<tr>
<td>ICON-7 CP</td>
<td>NR (22.9–NR)</td>
<td>14.9 (13.3–20.3)</td>
<td>8.3 (6.8–9.0)</td>
</tr>
<tr>
<td>ICON-7 CP + bevacizumab</td>
<td>20.8 (19.2–NR)</td>
<td>19.6 (15.9–20.9)</td>
<td>11.7 (9.6–13.5)</td>
</tr>
</tbody>
</table>

NOTE: KELIM AGO OVAR 9 Q33% = 0.0482 per days; Q66% = 0.0711 per days. Abbreviation: NR, not reached.

*Favorable tercile: KELIM in [Q66%–max]; intermediate tercile: KELIM in [Q33%–Q66%]; and unfavorable tercile: KELIM in [min–Q33%].

This European collaborative study aimed to confirm the previously reported predictive and prognostic values of the modeled kinetic CA-125 parameter KELIM in first-line treatment patients. The same mathematical model could be successfully used for characterizing the longitudinal CA-125 kinetics of patients enrolled in three large randomized clinical trials testing different chemotherapy regimens, thereby showing the feasibility of the approach. Moreover, this study demonstrated that the strength of the statistical association of survival with modeled CA-125 KELIM was much stronger than with the GCIG response criterion. The higher predictive value of KELIM regarding PFS and OS was confirmed. The magnitude of the discriminatory power of KELIM in terms of prognostic information is illustrated by the nonparametric and parametric survival tests. The patients with favorable CA-125 KELIM, above 0.0711 per days during the first 100 days, all experienced long median survivals regardless of their chemotherapy regimens, conversely to those with unfavorable KELIM. The benefit related to favorable KELIM was as high as approximately 50% consistent improvement in both PFS and OS, with clinically meaningful median survivals with respect to literature data (14, 28–32). Interestingly, the prognostic value of KELIM was maintained within ICON-7 risk groups reported by Oza and colleagues (14), who showed that adding bevacizumab to CP significantly improved the median PFS and OS among high-risk patients (14). On a practical point of view, individual KELIM values can be easily calculated online by any clinician on the dedicated internet site (http://www.biomarker-kinetics.org/).

Results from this study must be interpreted with caution. The semimechanistic model, used in this study, is a simplified version of the more complex models that were reported in previous publications, where CA-125 kinetics were linked to tumor burden changes (34) and to survivals (9). This type of mechanistic models would have helped to describe more accurately the exact relationships between tumor size, CA-125 production KPROD, and chemotherapy effect on KPROD, if the data required for estimating all the kinetic parameters of such complex models had been available in the three datasets. However they were not collected in the trials. As a consequence, we chose a simplified model to reduce the risk of biased estimations. Moreover this model was shown to provide reproducible and robust outcomes regarding CA-125 kinetic characterization and about the predictive value of the

Figure 3. OS probability versus time using univariate AFT model according to KELIM: categorized by the terciles (dotted line: unfavorable KELIM tercile [minimum–0.0482 per days]; dot-dash line: intermediate KELIM tercile [0.0482 per days–0.0711 per days]; long dash line: favorable KELIM tercile [0.0711 per days–maximum]) (A); or considered as a continuous covariate when KELIM/0.06 was arbitrarily set up at 0.5 (dotted line); 1 (dot-dash line); and 1.5 (long dash line). (B) 95% CIs on the AFT model were obtained using 1,000 simulations from the normal asymptotic distribution of the estimates. Black line, Kaplan-Meier without any covariate.

Accelerated failure time model

A

B

OS Time (months)

Survival probability (%)

www.aacrjournals.org
model kinetic parameter KELIM. Differences in trial designs and a lack of balance in enrolled patient characteristics, which were potentially increased by the selection criteria used for modeling, may have introduced heterogeneities in analyzed data and outcome biases (e.g., 85% of AGO OVAR 7 trial patients were excluded because of insufficient numbers of CA-125 values). However, the relative consistency of survival data from patients treated with the standard CP regimen across the three trials is reassuring. If the prognostic information provided by KELIM is strong, reproducible, and clinically relevant, the predictive power may be considered more limited, as shown by the C-index value close to 0.6 for OS. Although one could claim for better discrim-
natory predictive characteristics, modeled CA-125 kinetic KELIM exhibits similar predictive power to well-recognized prognostic factors of treatment efficacy, such as Federation Internationale des Gynaecologistes et Obstetrices (FIGO) disease stage, or Oza and colleagues risk groups in ICON-7 trial. It also displayed much stronger predictive value than the GCIG CA-125 response crite-
rion, which is still used in all ovarian cancer trials, despite the lack of any relevant prognostic or predictive information provided by this endpoint (5, 10). Of note, in this study, the GCIG CA-125 response criterion could not be calculated in about 5% of patients due to missing CA-125 baseline value.

Beyond the prognostic or predictive values of KELIM value itself, the surrogate value of treatment-related KELIM variations is of high interest for drug development. Indeed, if KELIM alterations induced by changes in treatment regimens were shown to be related to treatment efficacy effects, this endpoint could help select the best drug candidates during early-phase trials. Acknowledging the explorative aspect of this analysis due to the limited number of assessed trials, the linear regression suggested that KELIM alterations might explain more than 90% of observed OS differences, thereby highlighting KELIM changes might be used for predicting the survival benefit of novel treatments. In the ICON-7 trial high-risk group, an increase in the KELIM median value by 6.7% in the bevacizumab arm was associated with an observed 17.7% OS gain, consistently with reported findings (14). In theory, a 59% increase in KELIM would be required to obtain a 50% OS benefit with the addition of a new drug, provided the linear relationships between KELIM differences and OS changes were confirmed. If this assumption was validated, randomized phase II trials assessing the relative changes in KELIM could be used to select the best drug candidates or patient subgroups, by forecasting the potential survival benefits in subsequent phase III trials. This hypothesis will have to be confirmed using additional datasets from randomized trials, in a GCIG collaboration.

In summary, the data from three large phase III trials, along with those previously reported in the CALYPSO trial (more than 3,500 patients), all suggest that the strength of the statistical association of survival with the modeled parameter KELIM is stronger than with CA-125 GCIG response criterion as far as CA-125 longitudinal kinetics is concerned. KELIM, easily calculable online, is a candidate endpoint for replacing GCIG CA-125 response criterion in international guidelines.

Disclosure of Potential Conflicts of Interest

J. Pfisterer is a consultant/advisory board member for Roche, Amgen, AstraZeneca, MSD Sharp & Dohme GmbH, Clovis, and Tesaro. A. Du Bois is a consultant/advisory board member for AstraZeneca, Roche, Tesaro, Clovis, BiOCAD, Pfizer, and PharmAmar. C. Kurzeder holds ownership interest (including patents) in AstraZeneca and PharmAmar, is a consultant/advisory board member for Roche, AstraZeneca, Tesaro, Seattle Genetics, Pfizer, and Genomic Health, and has provided expert testimony for AstraZeneca with respect to olaparib in first-line treatment of ovarian cancer. No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions

Conception and design: O. Colomban, J. Pfisterer, A. Du Bois, C. Kurzeder, B. You

Development of methodology: O. Colomban, M. Tod, A. Burges, J. Péron, B. You

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): J. Ray-Coquard, A. Lortholary, A.C. Hardy-Bessard, J. Pfisterer, A. Du Bois, C. Kurzeder, A. Burges, G. Freyer, B. You


Writing, review, and/or revision of the manuscript: O. Colomban, M. Tod, A. Leary, J. Ray-Coquard, A. Lortholary, J. Pfisterer, A. Du Bois, B. You

Study supervision: A. Lortholary, J. Pfisterer, A. Du Bois, B. You

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): A. Lortholary, A.C. Hardy-Bessard, A. Du Bois, J. Péron, G. Freyer, B. You

Acknowledgments

The authors thank all the patients and their families, the investigators, the study nurses, pathologists, and all study teams, especially the Gynecologic Cancer Intergroup and MRC Clinical Trial Unit at University College London. This work was supported by Université Claude Bernard Lyon 1, Association de Recherche sur les Cancers dont Gynécologiques (ARCAGY)-Groupe des Investiateurs Nationaux des Cancers de l’Ovarie (GINECO), Arbeitsgemeinschaft Gynäkologische Onkologie (AGO), Gynecologic Cancer Intergroup (GCIG), and MRC Clinical Trial Unit at University College London.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received October 12, 2018; revised January 3, 2019; accepted March 28, 2019; published first April 1, 2019.

References


8. Chiang AI, Chen JR, Chung YH, Huang HJ, Liu WS, Chang C. A longi-
25. Jonsson EN, Karlsson MO. Xpose—an S-PLUS based population pharmaco-
# Early Modeled Longitudinal CA-125 Kinetics and Survival of Ovarian Cancer Patients: A GINECO AGO MRC CTU Study

Olivier Colomban, Michel Tod, Alexandra Leary, et al.

*Clin Cancer Res*  Published OnlineFirst April 1, 2019.

## Updated version
Access the most recent version of this article at: doi:10.1158/1078-0432.CCR-18-3335

## Supplementary Material
Access the most recent supplemental material at:
- [http://clincancerres.aacrjournals.org/content/suppl/2019/03/30/1078-0432.CCR-18-3335.DC1](http://clincancerres.aacrjournals.org/content/suppl/2019/03/30/1078-0432.CCR-18-3335.DC1)
- [http://clincancerres.aacrjournals.org/content/suppl/2019/05/09/1078-0432.CCR-18-3335.DC2](http://clincancerres.aacrjournals.org/content/suppl/2019/05/09/1078-0432.CCR-18-3335.DC2)

## E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

## Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

## Permissions
To request permission to re-use all or part of this article, use this link [http://clincancerres.aacrjournals.org/content/early/2019/05/03/1078-0432.CCR-18-3335](http://clincancerres.aacrjournals.org/content/early/2019/05/03/1078-0432.CCR-18-3335). Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.