

Integrative Protein-Based Prognostic Model for Early-Stage Endometrioid Endometrial Cancer

Ji-Yeon Yang^{1,2}, Henrica M.J. Werner^{3,4}, Jie Li⁵, Shannon N. Westin⁶, Yiling Lu⁵, Mari K. Halle^{3,4}, Jone Trovik^{3,4}, Helga B. Salvesen^{3,4}, Gordon B. Mills⁵, and Han Liang^{1,5}

Abstract

Purpose: Endometrioid endometrial carcinoma (EEC) is the major histologic type of endometrial cancer, the most prevalent gynecologic malignancy in the United States. EEC recurrence or metastasis is associated with a poor prognosis. Early-stage EEC is generally curable, but a subset has high risk of recurrence or metastasis. Prognosis estimation for early-stage EEC mainly relies on clinicopathologic characteristics, but is unreliable. We aimed to identify patients with high-risk early-stage EEC who are most likely to benefit from more extensive surgery and adjuvant therapy by building a prognostic model that integrates clinical variables and protein markers.

Experimental Design: We used two large, independent early-stage EEC datasets as training ($n = 183$) and validation cohorts ($n = 333$), and generated the levels of 186 proteins and phosphoproteins using reverse-phase protein arrays. By applying an initial filtering and the elastic net to the training samples, we

developed a prognostic model for overall survival containing two clinical variables and 18 protein markers and optimized the risk group classification.

Results: The Kaplan–Meier survival analyses in the validation cohort confirmed an improved discriminating power of our prognostic model for patients with early-stage EEC over key clinical variables (log-rank test, $P = 0.565$ for disease stage, 0.567 for tumor grade, and 1.3×10^{-4} for the integrative model). Compared with clinical variables (stage, grade, and patient age), only the risk groups defined by the integrative model were consistently significant in both univariate and multivariate analyses across both cohorts.

Conclusions: Our prognostic model is potentially of high clinical value for stratifying patients with early-stage EEC and improving their treatment strategies. *Clin Cancer Res*; 22(2); 513–23. ©2015 AACR.

Introduction

Endometrial cancer is the most prevalent gynecologic malignancy in the United States, with an estimated 54,870 new cases and 10,170 deaths occurring in 2015 (1). In contrast to progress observed for most cancer types, the incidence and mortality from endometrial cancer has increased over the past several decades (2). The most common histologic type of endometrial cancer

(~80%), endometrioid endometrial carcinoma (EEC), is linked to estrogen excess, obesity, and hormone-receptor positivity. Nonendometrioid tumors include serous carcinoma (~10%) and several other rare types (e.g., clear cell carcinomas and carcinosarcomas; ref. 3), which are typically associated with advanced stage at diagnosis and poor prognosis. These tumors are usually treated with more aggressive therapy, including total abdominal hysterectomy, bilateral salpingo-oophorectomy, and postsurgical radiation therapy and chemotherapy (4, 5).

The reported 5-year survival rate is relatively good in early-stage EEC (stage I, II, 70%–90%) but it falls below 50% in late-stage EEC (40%–50% for stage III, 15%–20% for stage IV; ref. 6). Indeed, patients with late-stage EEC or recurrent disease have a poor prognosis, with a median overall survival of less than 12 months. Their treatment options are similar to those offered to patients with nonendometrioid carcinoma (7). Fortunately, most women with EEC are diagnosed at an early stage of the disease, when the cancer can generally be cured by surgery. However, not all early-stage EEC is clinically indolent: a subset has poor prognosis due to metastasis or disease recurrence following surgery (8). This clinical heterogeneity for early-stage EEC warrants efforts to identify high-quality prognostic markers that can be used to stratify patients and to determine their prognoses, thereby implementing evidence-based individualized therapeutic approaches. Currently, the prediction of prognosis for patients with early-stage EEC is mainly based on disease stage and tumor grade. Other prognostic factors include lymphovascular space invasion, obesity, age, myometrial invasion depth, and lymph node status

¹Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas. ²Department of Applied Mathematics, Kumoh National Institute of Technology, Gumi-si, South Korea. ³Centre for Cancer Biomarkers, Department of Clinical Science, The University of Bergen, Bergen, Norway. ⁴Department of Obstetrics and Gynecology, Haukeland University Hospital, Bergen, Norway. ⁵Department of Systems Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas. ⁶Department of Gynecologic Oncology and Reproductive Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (<http://clincancerres.aacrjournals.org/>).

Corresponding Authors: Han Liang, Department of Bioinformatics and Computational Biology, Unit 1410, The University of Texas MD Anderson Cancer Center, 1400 Pressler Street, Houston, TX 77030. Phone: 713-745-9815; Fax: 713-563-4242; E-mail: hliang1@mdanderson.org; Gordon B. Mills, Department of Systems Biology, Unit 950, 1515 Holcombe Boulevard, Houston, TX 77030. E-mail: gmills@mdanderson.org

doi: 10.1158/1078-0432.CCR-15-0104

©2015 American Association for Cancer Research.

Translational Relevance

A subset of endometrioid endometrial carcinoma (EEC) is associated with recurrence or metastasis, leading to a poor prognosis. Currently, prognosis estimation for early-stage EEC mainly relies on clinicopathologic characteristics, but is unreliable. Here, we have developed and validated an effective postoperative prognostic model for the overall survival of patients with early-stage EEC, which integrates clinical variables and protein markers. To the best of our knowledge, this is the first quantitative protein-based model that represents a novel, improved prognostic approach for this patient population. The risk score calculated from this model will help identify patients with high-risk early-stage EEC who are most likely to benefit from more extensive surgery and adjuvant therapy.

(9, 10), but none of these factors can reliably predict patient survival. Therefore, it is critical to develop an effective strategy to identify patients with high-risk, early-stage EEC because they are likely to benefit from more extensive surgery and aggressive adjuvant therapies. Furthermore, this may spare patients who have low-risk EEC from potentially toxic interventions.

The detailed characterization of endometrial tumors at the molecular level may provide information that can be used to increase the power of prognostic models beyond that offered by clinicopathologic characteristics alone, thereby improving surveillance and treatment strategies. In this study, we aimed to develop an effective prognostic model for patients with early-stage EEC by integrating clinical characteristics with proteomic data. As proteins are major functional units in various biologic processes, protein markers have the potential to reflect key processes underlying tumor development and progression that may not be captured by other types of molecular markers (11, 12). To identify and validate potential biomarkers, we used functional proteome-based reverse-phase protein array (RPPA), which is an antibody-based technology that allows for high-throughput measurements of the expression levels of proteins and phosphoproteins in a large number of samples in a cost-effective and sensitive manner (13, 14). The technical reliability and the utility of this platform have been well documented in the literature, including in our previous studies (15–17). A typical coefficient of variation for RPPA is 5% to 7%, demonstrating high reproducibility (15, 18). In our recent pan-cancer study, the correlation analysis among RPPA protein markers successfully captured known signaling pathways in different tumor contexts (11). Importantly, through a systematic assessment on prognostic utility, we found that RPPA-based protein expression is the most informative data among various types of molecular data surveyed (i.e., somatic copy-number alterations, DNA methylation, mRNA, microRNA and protein expression; ref. 19). Therefore, we focused on the RPPA-based protein markers in this study.

Here, we used two large, independent patient sets as the training and validation cohorts, respectively. On the basis of their survival data availability/quality and also because we are more interested in patients' long-term outcome, we considered overall survival as the endpoint in our analysis. We first built a prognostic model for patients with early-stage EEC from the training cohort,

and determined a risk group classification scheme. Our prognostic model that integrates both protein markers and clinical characteristics showed better performance than conventional clinical variables in both the training and validation cohorts. Thus, the integrative prognostic model we have developed may represent a valuable tool for improving the surveillance and treatment strategies of patients with early-stage EEC.

Materials and Methods

Patient sample collection

In this study, fresh-frozen samples were collected from patients newly diagnosed with EEC at Haukeland University Hospital (Bergen, Norway), the University of Texas MD Anderson Cancer Center (MDACC; Houston, Texas), and The Cancer Genome Atlas (TCGA). Bergen samples were approved by the Norwegian Data Inspectorate (961478–2), the Norwegian Social Science Data Services (15501), and the local Institutional Review Board (REKIII nr. 052.01). The MDACC samples were approved by the Institutional Review Board of MDACC (Lab08–0580). Treatment plans followed the established standards at the respective institutions in accordance with NCCN guidelines and as previously reported (20). In general, early-stage EEC was treated with surgery alone with or without adjuvant radiotherapy; late-stage EEC was treated with surgery followed by chemotherapy with or without volume-directed radiotherapy. Tumor content, histologic classification, grade, and disease stage were reviewed by independent pathologists. All the patients provided written informed consent for the collection of samples and subsequent analyses.

For prognostic modeling, we used 209 EEC samples from the Bergen set as the training cohort because of the large sample size and long follow-up time associated with this dataset. We combined the MDACC and TCGA samples as the validation cohort ($n = 427$) to increase the total sample size and statistical power. The clinical information for TCGA samples was obtained from the TCGA consortium paper for endometrial cancer (21). We had limited clinical data, especially for MDACC and Bergen samples. Among potential affecting factors, patient age, grade, disease stage, body mass index (BMI), and ethnicity were available for MDACC samples, and patient age, grade and disease stage for Bergen samples. Therefore, we used three clinical factors (patient age, disease stage, and tumor grade) that were available for all patients in the three sample sets. We capped the overall survival time at 60 months, as described previously (22). The individual clinical data of the patients included in this study are reported in Supplementary Tables S1 and S2.

Reverse-phase protein array profiling and data normalization

We performed RPPA profiling for all samples at the MD Anderson RPPA core facility, as previously described (21). Briefly, proteins were extracted from tumor tissues, denatured by SDS, and printed on nitrocellulose-coated slides, which was followed by antibody probing. Quantification of protein expressions was performed through "supercurve fitting," as described previously (23, 24), and which is a method that has been extensively validated for both cell line and patient samples (15, 17). We obtained quantitative protein expression profiles of 186 proteins and phosphoproteins for both the training and validation samples (Supplementary Table S3 for the list of 186 antibodies). These

proteins were selected to broadly represent the important pathways in cancer, and their RPPA data have been widely used in various TCGA-related analyses, providing deep insights into the molecular mechanisms of various cancer types. All the antibodies were validated by Western blot analysis. The validation process and assessment results on the antibodies for RPPA were detailed previously (15).

All the samples used in this study were run in three different RPPA batches. Because there is a potential problem of batch effects when combining the batches of protein expression data, we used replicate-based normalization (RBN) method as described in our recent study (11), which uses replicate samples run across multiple batches. In one "anchor" batch, we ran many replicate samples that were common with the other two batches (200 and 30, respectively). Those samples were all TCGA validation samples and non-control samples. RBN was performed by adjusting each data point in the non-anchor batches so that the mean and variance of the common samples for each antibody are identical to those in the anchor batch. The normalized RPPA data for the 209 training samples and 427 validation samples are provided in Supplementary Tables S1 and S2.

Development of the prognostic model for early-stage EEC

We used data from 183 patients with stage I or II EEC from the Bergen set as the training cohort. As a preprocessing step to reduce feature dimensionality, we filtered out noisy features by applying univariate Cox regressions to protein expression and clinical features (186 protein markers and 3 clinical factors). To avoid ruling out potentially relevant features, we used a cutoff of 0.15 for the *P* values, and retained only features with a *P* value smaller than 0.15 for model development.

Then, we used the elastic net (25) to identify markers associated with overall survival and to use in training the final model for prediction with the selected features. The elastic net simultaneously conducts automatic variable selection and group selection of the correlated variables. The explicit objective function and the algorithm for estimating the solution of the elastic net has been described previously (25, 26). We used the R package "glmnet" for the implementation (27). We used leave-one-out cross-validation to select the tuning parameter, and determined the elastic net mixing parameter in order to find a parsimonious model while maintaining a modest discriminating accuracy based on the concordance index (28).

The final model is a linear combination of features selected by the elastic net, weighted by the corresponding elastic net coefficients. The weights are a rough estimate of the contribution of the information content of each marker to the overall risk score. Specifically,

$$\begin{aligned} \text{Risk score} = & 0.048 \times \text{age} + 0.118 \times \text{grade} - 1.032 \times \text{EGFR} \\ & - 0.659 \times \text{phospho-MyosinIIa} - 0.496 \times \text{AR} - 0.462 \times \text{cMyc} \\ & - 0.273 \times \text{phospho-STAT3} - 0.131 \times \text{phospho-p38 MAPK} \\ & - 0.065 \times \text{phospho-mTOR} - 0.055 \times \text{fibronectin} - 0.018 \\ & \times \text{Hsp70} + 0.071 \times \text{FASN} + 0.179 \times \text{acetyl-}\alpha\text{-Tubulin (Lys40)} \\ & + 0.210 \times \text{ACC1} + 0.249 \times \text{Ku80} + 0.418 \times \text{phospho NF} \\ & \kappa\text{B p65} + 0.420 \times \text{Bim} + 0.504 \times \text{Chk2} + 0.573 \times \text{cKIT} \\ & + 1.527 \times \text{SMAD4} \end{aligned} \quad (1)$$

Because the model includes both protein markers and clinical variables from samples of early-stage EEC, we refer to the model as

the integrative prognostic model for early-stage EEC. The exponential of an elastic net coefficient gives the hazard ratio (HR) of death associated with each marker. However, there is no consensus on a statistically valid method of estimating the standard error of a coefficient estimate with shrinkage methods including the elastic net, so the standard errors or confidence intervals for elastic net estimates were not reported here.

To facilitate clinical application, we divided the training samples into two risk groups according to each patient's risk score based on Equation (1). We determined the cutoff value for the risk scores to ensure that the two risk groups would have similar numbers of events, as described previously (29). Statistically, it is known that analysis of groups with unbalanced numbers of events may introduce a bias in the parameter estimation and the bias will be substantial when the number of events is small, that is, in case of rare events (30, 31). In EEC (especially in early-stage EEC), mortality is not so high and thus the number of events is relatively small. Therefore, we used a simple but widely used approach of splitting samples into risk groups based on equal/similar number of events, in order to reduce the estimation bias and to ensure the similar standard errors of the parameter estimates across risk groups.

Validation of the integrative prognostic model for early-stage EEC

We used 333 samples of stage I and II EEC from the validation cohort to validate the prognostic model [Equation (1)]. We computed risk scores based on Equation (1) for the validation samples, and then classified them into low- or high-risk groups, with the cutoff determined in the training set as described above. We used univariate and multivariate Cox regressions to evaluate the patient risk classification.

Development and validation of the prognostic model for late-stage EEC

In a similar way, we constructed and validated a prognostic model for patients with late-stage EEC. We trained the model using samples of stages III and IV EEC from the Bergen set ($n = 26$) by initially filtering the data through a univariate Cox regression, with a *P* value of 0.15 as the cutoff, and selecting the features through the elastic net, with the tuning parameter selected by leave-one-out cross-validation. The final model is as follows:

$$\begin{aligned} \text{Risk score} = & 0.019 \times \text{age} + 0.203 \times \text{grade} - 1.511 \times \text{SMAD4} \\ & - 0.479 \times \text{phospho-Shc} - 0.421 \times \text{cKIT} - 0.270 \times \text{phospho} \\ & \text{-ERK/MAPK} - 0.207 \times \text{PKC}\alpha - 0.180 \times \text{Bim} - 0.099 \times \text{Bcl2} \\ & + 0.113 \times \text{TFRC} + 0.179 \times \text{ASNS} + 0.336 \times \text{VEGFR-2} \\ & + 0.364 \times \text{PAI1} + 0.392 \times \text{phospho-HER2} + 0.446 \\ & \times \text{phospho-4E-BP1} + 1.311 \times \text{EGFR} \end{aligned} \quad (2)$$

This model also contains both protein markers and clinical features, so we refer to it as the integrative prognostic model for late-stage EEC. We determined the risk score cutoff for the classification of the training samples such that the two risk groups would have similar numbers of events.

We computed risk scores for the samples of late-stage EEC (stages III and IV, $n = 94$) in the validation cohort, and divided the samples into the two risk groups according to the same cutoff used for the training samples.

Comparison with clinical-variable-only models

For the practical utility of the integrative models, we considered the patient's classification as being in the low- or high-risk group instead of risk scores. To evaluate the performance of our integrative models relative to prognoses based on only clinical variables, we considered univariate and multivariate Cox proportional hazards models, with the following characteristics as covariates:

1. Disease stage
2. Tumor grade
3. Patient age
4. Risk group index based on the integrative model
5. Stage + grade + age + risk group index based on the integrative model

For the disease stage, we compared two levels: stage I vs. stage II in the early-stage EEC model and stage III vs. stage IV in the late-stage EEC model. In the Cox regressions, we treated the tumor grade as a continuous variable, with natural ordering so as to compute only one regression coefficient for this variable and to examine the overall difference across tumor grades. We dichotomized the patient's age as two categories, younger than 60 years of age or 60 years of age and older, in order to evaluate the efficacy of a clinical practice guideline: patients 60 years of age and older are generally advised to receive adjuvant therapy (32). Furthermore, based on a previous study (10), we considered three categories of patient age (<50, [50, 70), and ≥ 70), but found no events in the group of <50 years. Instead, we considered an alternative, similar three-age classification (<60, [60, 80) and ≥ 80). In addition, we computed concordance indexes (C-indexes) to compare the discriminatory power of clinical factors and the integrative models.

We evaluated the additional prognostic value of the integrative models over the clinical factors using a multivariate regression analysis with Model 5: Stage + Grade + Age + Risk group index based on the integrative model. Although our integrative models include the two clinical factors of patient age and tumor grade,

multicollinearity is not an issue in Model 5 because we used the risk group index rather than the risk score as a covariate. In addition, we used log-rank tests to examine the differences in survival between the risk groups as stratified by the integrative models, between disease stages and among tumor grades. In order to assess the robustness of our models to different cutoff values, we tried different cutoffs that made the sample numbers in risk groups equal to those numbers with different stages (or different grades). We also considered another different cutoff value, the 25th percentile of risk scores (which approximately corresponded to the value of 0.5 in early-stage samples and 2.5 in late-stage samples).

Results

Patient characteristics and the prognostic power of disease stage and tumor grade

For robust prognostic modeling, we used two independent datasets of EEC as training and validation cohorts. The patients' characteristics are summarized in Table 1. Our training cohort was obtained from Haukeland University Hospital (Supplementary Table S1), and contained 183 samples of early-stage EEC (FIGO 2009 stages I and II) and 26 samples of late-stage EEC (stages III and IV). The validation samples were obtained from MDACC and TCGA (Supplementary Table S2). As there was no significant difference in survival between the MDACC and TCGA datasets for either early-stage or late-stage EEC (log-rank test, $P = 0.36$ for early-stage EEC and 0.85 for late-stage EEC), we combined them as one validation cohort to increase the sample size and boost the statistical power. In total, the validation cohort contained samples from 333 patients with early-stage EEC and 94 patients with late-stage EEC. We generated the expression profiles of 186 proteins and phosphoproteins using RPPA. The RPPA data for the training and validation samples are presented in Supplementary Tables S1 and S2, and information about the 186 antibodies is provided in Supplementary Table S3. The proteomic profiling and quality

Table 1. Clinical characteristics of the training and validation cohorts

	Training		Validation		P^a (1) vs. (3)	P^b (2) vs. (4)
	Early-stage EEC (1)	Late-stage EEC (2)	Early-stage EEC (3)	Late-stage EEC (4)		
# of patients	183	26	333	94		
Age: mean (range)	64.94 (35-93)	67.92 (43-89)	61.38 (24-90)	62.06 (32-89)	1.4×10^{-3}	0.029
Disease stage ^c : N (%)					0.22	0.69
I	170 (92.9)		297 (89.2)			
II	13 (7.1)		36 (10.8)			
III		18 (69.2)		71 (75.5)		
IV		8 (30.8)		23 (24.5)		
Tumor grade: N (%)					5.3×10^{-4}	0.80
Grade 1	74 (40.4)	3 (11.5)	84 (25.2)	7 (7.5)		
Grade 2	82 (44.8)	11 (42.3)	167 (50.2)	41 (43.6)		
Grade 3	27 (14.8)	12 (46.2)	82 (24.6)	46 (48.9)		
Deaths ^d : N (%)	17 (9.3)	16 (61.5)	21 (6.3)	20 (21.3)		
Recurrences ^e : N (%)	27 (14.8)	18 (69.2)	33 (10.8)	43 (49.4)		

^a P values comparing the patients with early-stage EEC in the training cohort with those in the validation cohort. For age, the two-sample t test was used. For stage and grade, the χ^2 test was used.

^b P values comparing the patients with late-stage EEC in the training cohort with those in the validation cohort. For age, a two-sample t test was used. For stage and grade, the χ^2 test was used.

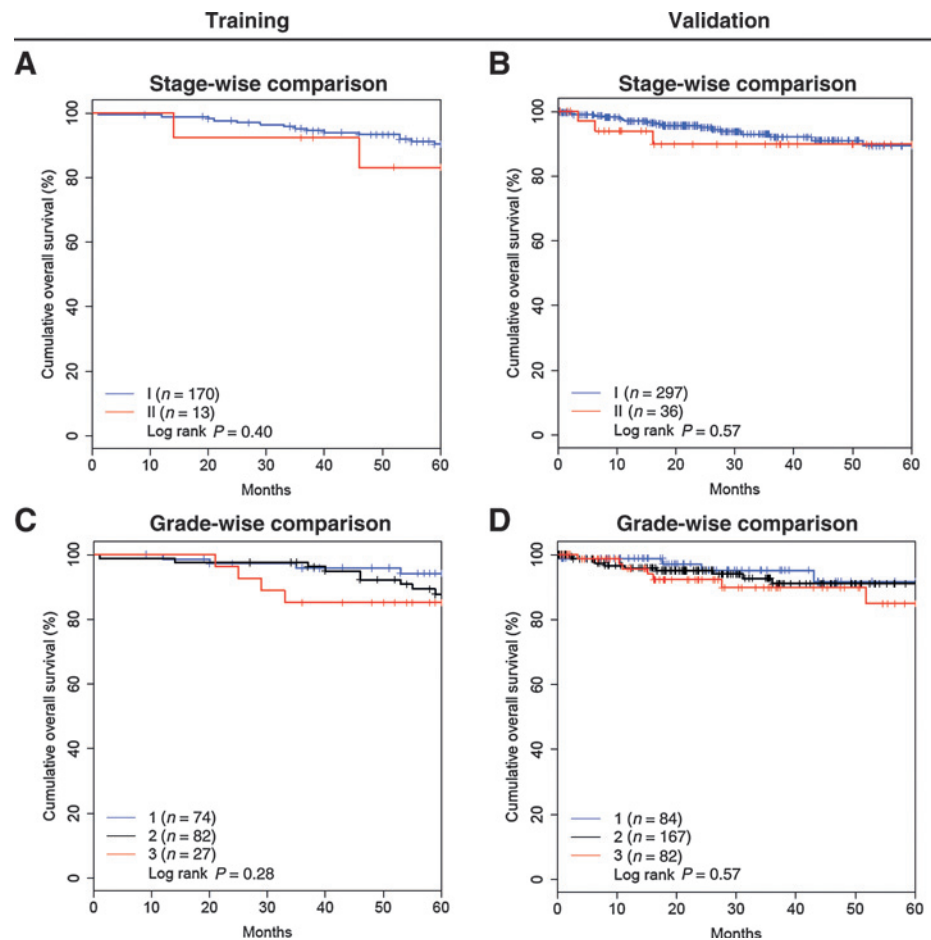
^cFIGO 2009 Criteria.

^dSurvival times were capped at 60 months.

^eData on recurrence were not available for 27 early-stage validation samples and 7 late-stage validation samples. Recurrence time was defined as the time from initial surgery to the first documented progression or recurrence or the last follow-up in the absence of progressive disease. Recurrence times were capped at 60 months.

Figure 1.

The Kaplan-Meier survival curves for patients with early-stage EEC (stages I and II) in the training and validation cohorts. For A and B, patients are split by disease stage; for C and D, patients are split by tumor grade.



control followed the well-established procedures in the TCGA project (12, 21). To remove batch effects in RPPA data, we used replicate-based normalization, as previously described (11).

Patients represented in the training cohort were slightly older than those in the validation cohort (early-stage EEC: mean age in training 64.9 years vs. mean age in validation 61.4 years, *t* test, $P = 1.4 \times 10^{-3}$; late-stage EEC: mean age in training 67.9 years vs. mean age in validation 62.1 years, $P = 0.029$; Table 1). No significant difference in the stage distributions was found between the training and validation cohorts, using χ^2 tests ($P = 0.22$ for early-stage EEC and 0.69 for late-stage EEC). We found a significant difference in the grade distributions between the training and validation cohorts for the patients with early-stage EEC ($P = 5.3 \times 10^{-4}$): samples of early-stage EEC in the validation cohort contained a larger proportion of high-grade tumors, which may reflect the strategy to enrich for large tumors in the TCGA compared with the population-based approach for the Bergen cohort (33). There was no statistically significant difference in the overall survival time between the training and validation cohorts for the patients with early-stage EEC (log-rank test, $P = 0.48$); however, the patients with late-stage EEC who were represented in the training cohort showed a significantly worse survival time than those in the validation cohort (log-rank test, $P = 0.014$; Supplementary Fig. S1).

We evaluated the discriminating power of disease stage and tumor grade for the early-stage and late-stage EEC samples. For the

early-stage EEC samples, there was no difference in overall survival when the patients were stratified by disease stage or tumor grade: log-rank test *P* values were 0.40 and 0.57 for the respective training and validation cohorts when split by disease stage (Fig. 1A and B; the results also held true when substages were considered); and the log-rank test *P* values were 0.28 and 0.57 for the respective training and validation cohorts when split by tumor grade (Fig. 1C and D).

For the late-stage EEC samples, there was a significant survival difference in the training cohort when split by disease stage (log-rank test, $P = 0.02$) and in the validation cohort when split by tumor grade (log-rank test, $P = 0.02$). Supplementary Fig. S2 shows the Kaplan-Meier curves for late-stage samples when patients were stratified according to disease stage or tumor grade. Thus, no significant discriminating power of stage or grade was observed for early-stage EEC, which highlights a pressing need for effective prognostic models for this patient population. On the basis of the markedly different marker patterns and clinical outcomes, we performed prognostic modeling on early-stage EEC samples and late-stage EEC samples separately.

An effective prognostic model for patients with early-stage EEC

The flow chart in Fig. 2 shows the overall procedure of developing and validating an integrative prognostic model for patients with early-stage EEC. We developed a prognostic model after applying an initial filtering and the elastic net to the early-stage

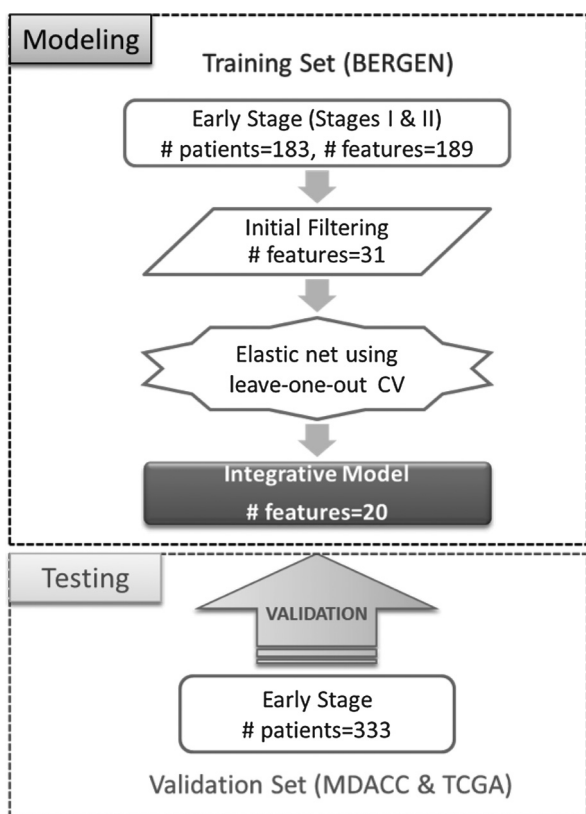


Figure 2. A flow chart for construction and validation of the prognostic model for overall survival in early-stage endometrioid endometrial cancer.

samples in the training cohort (see Materials and Methods). The integrative model [Equation (1) in Materials and Methods] contains two clinical factors (patient age and tumor grade) and 18 protein markers (listed in Supplementary Table S4). On the basis of this model, we computed risk scores for the training samples as a weighted sum of the selected features. As shown in Supplementary Fig. S3A, the distribution of risk scores was unimodal with very similar peaks between early-stage training and validation samples. To facilitate clinical application, we chose a risk score cutoff to classify patients with early-stage EEC in the training cohort into low- and high-risk groups. The cutoff was determined to ensure that the two risk groups would have similar numbers of events; and this scheme resulted in a cutoff point of approximately the 90th percentile of risk scores in the training cohort, as illustrated in the second column of Supplementary Fig. S3A. As expected, survival was much worse among patients in the high-risk group compared with those in the low-risk group (log-rank test, $P = 2.3 \times 10^{-14}$; Fig. 3A). From the clinical point of view, this risk score cutoff point is of potential utility because it helps identify a subset of patients with early-stage EEC (10%) who are most likely to benefit from additional treatment despite added toxicity.

To evaluate the discriminating power of our integrative model, we *locked* the model and applied it to samples of early-stage EEC in the validation cohort. The training and validation samples showed similar ranges and distributions of risk scores. Using the same cutoff value, we split the validation samples into two risk groups with similar numbers of events as described above. The

Kaplan–Meier survival analysis showed a significant survival difference between the two risk groups (log-rank test, $P = 1.3 \times 10^{-4}$; Fig. 3B).

We used a data-driven cutoff (i.e., the 90th percentile of risk scores of the training samples) as described previously (29, 34); the determination of this cutoff point is a part of our modeling process. To assess the robustness of our model to different cutoff values, we considered cutoffs that gave equal sample sizes based on the same disease stage (or tumor grade). In comparison with the power of the disease stage (stage I vs. II), the 183 early-stage EEC training samples were split into two risk groups, with $n = 170$ and $n = 13$, which were the numbers of training samples in stage I and II, respectively. Similarly, the 297 validation samples with the lowest risk scores were classified as low risk and the remaining 36 validation samples were classified as high risk. We observed a significant survival difference between these two risk groups (log-rank test, $P = 1.0 \times 10^{-14}$ for the training samples; and $P = 2.5 \times 10^{-3}$ for the validation samples; Fig. 3C and D). When using cutoffs that gave the same sample sizes based on the tumor grades (grades 1, 2, and 3), we also observed a statistically (or marginally) significant difference in overall survival among the risk groups (log-rank test, $P = 4.0 \times 10^{-10}$ for the training samples; and $P = 0.066$ for the validation samples, Fig. 3E and F). Furthermore, using the 25th percentile of risk scores, which was approximately equal to 2.5, there was statistically significant (or marginal significance) discrimination of patients with regard to survival in early-stage training and validation samples, as shown in Supplementary Fig S4A and S4B (log-rank test, P values are 0.01 and 0.076). These separations, however, were not as distinct as those for which the cutoff was driven by the data or the cutoff ensured equal sample sizes for the different stages. In addition, we considered patient stratification according to the median split, and observed significantly improved survival in the low-risk group for both training and validation samples. Taken together, these results indicate that patient stratification based on our integrative model is effective and robust across multiple cutoff points.

We further investigated the additional discriminating power of this integrative model over the clinical factors using univariate and multivariate Cox proportional hazards models. We considered three clinical factors (disease stage, tumor grade, and patient age) that were available for both the training and validation cohorts in the analysis. Age was dichotomized at a threshold of 60 years because adjuvant therapy is generally recommended for patients who are 60 years of age or older (32). Age was significant in the univariate model for the validation samples (Wald test, $P = 0.044$), but was neither significant in the univariate model for the training samples, nor in the multivariate models. In a line with a previous study (10), we also examined the effect of a three-age-group classification, and obtained very similar results if age was categorized as <60 , $[60, 80)$, and ≥ 80 years. The risk group index (based on our integrative model) was the only factor consistently significant in the univariate and multivariate models across the different datasets (Table 2). The integrative model showed the highest value of C-index when compared with the other clinical factors as presented in Table 2.

Prognostic modeling for patients with late-stage EEC

In parallel, we performed prognostic modeling for patients with late-stage EEC; however, (i) a relatively small sample size may limit our capacity to obtain a reliable model, and (ii) it may

Yang et al.

Table 2. Univariate and multivariate Cox proportional hazards model analysis for the patients with early-stage EEC

	Univariate analysis			
	HR ^a (90% CI)	Wald P ^b	Log-rank P ^c	C-Index
Training set				
Disease stage: II (vs. I)	1.88 (0.55, 6.47)	0.40	0.40	0.53
Tumor grade ^d	1.71 (0.98, 2.98)	0.12	0.28	0.60
Patient age ^e : ≥60 y (vs. <60 y)	2.38 (0.84, 6.76)	0.17	0.16	0.57
Integrative model: high risk (vs. low risk)	16.72 (7.39, 37.80)	1.5 × 10⁻⁸	2.3 × 10⁻¹⁴	0.76
Validation set				
Disease stage: II (vs. I)	1.43 (0.51, 3.98)	0.57	0.57	0.54
Tumor grade ^d	1.38 (0.83, 2.30)	0.30	0.57	0.56
Patient age ^e : ≥60 y (vs. <60 y)	3.05 (1.23, 7.61)	0.044	0.034	0.60
Integrative model: high risk (vs. low risk)	5.26 (2.37, 11.65)	6.3 × 10⁻⁴	1.3 × 10⁻⁴	0.70
Multivariate analysis^f				
	HR ^a (90% CI)	Wald P ^b		
Training set				
Disease stage: II (vs. I)	1.29 (0.37, 4.50)	0.74		
Tumor grade ^d	1.31 (0.71, 2.39)	0.47		
Patient age ^e : ≥60 y (vs. <60 y)	0.83 (0.24, 2.94)	0.81		
Integrative model: high risk (vs. low risk)	16.59 (6.08, 45.22)	4.4 × 10⁻⁶		
Validation set				
Disease stage: II (vs. I)	1.87 (0.66, 5.29)	0.33		
Tumor grade ^d	1.04 (0.60, 1.79)	0.91		
Patient age ^e : ≥60 y (vs. <60 y)	2.50 (0.98, 6.39)	0.11		
Integrative model: high risk (vs. low risk)	4.35 (1.82, 10.37)	5.6 × 10⁻³		

NOTE: P values < 0.1 are shown in bold.

^aHR: hazard ratio.^bWald test P values.^cLog-rank test P values.^dIn Cox regressions, grade was treated as a continuous variable with natural ordering so as to compute only one regression coefficient for this variable. However, log-rank test P values were computed by treating grade as a discrete variable.^eAge was dichotomized to 60 years and older vs. younger than 60 years because patients 60 years of age and older are generally advised to receive adjuvant therapy.^fOn the basis of a multivariate Cox regression model that included all the variables in the table.

training and validation samples, as shown in Supplementary Fig. S3B. To classify low- and high-risk groups, we used the 75th percentile of the risk scores in the training samples as a risk score cutoff to ensure a similar number of events in each risk group, as shown in the second column of Supplementary Fig. S3B.

We next applied the model to samples of late-stage EEC in the validation cohort and classified them into two risk groups using the 75th percentile as a cutoff value. The Kaplan–Meier curves for the two risk groups showed a clearer separation in survival as in Supplementary Fig. S5B than when the patients were stratified by disease stage as in Supplementary Fig. S2B; however, the separation in survival was not as clear as when patients were stratified by tumor grade as in Supplementary Fig. S2D regardless of the P values. We observed similar patterns with varying cutoff points while ensuring equal sample sizes between the risk groups split by disease stage or tumor grade, as we did for the analysis of early-stage EEC. The Kaplan–Meier curves for validation samples are displayed in Supplementary Fig. S5D and S5F, when patients were grouped by the sample numbers of disease stage or tumor grade. Furthermore, we observed similar results using the 25th percentile of risk scores, which was approximately equal to 0.5 (Supplementary Fig. S4C and S4D). Overall, among patients with late-stage EEC, the model based on only the tumor grade outperformed the integrative model in terms of prognostic accuracy. In multivariate analysis for the validation samples, we found that tumor grade, patient age, and the risk group index defined by the integrative model were all significant (Supplementary Table S6). Thus, the additional prognostic value obtained from the late-stage EEC integrative model compared with using clinical factors alone

seemed modest, and was mainly due to the small number of training samples and intrinsic differences in survival between the patients represented by the training and validation samples, as observed in Supplementary Fig. S1B.

Discussion

In this study, we aimed to improve prognosis estimation of EEC patients by incorporating high-throughput protein expression data. We did not intend to invent a new and alternative strategy completely independent from the current staging system. Rather, we used the clinical stage as an initial stratification parameter (by classifying patients into early vs. advanced stages: I, II vs. III, IV) and further incorporated protein-based markers designed to provide informative content that the disease stage may not capture. Therefore, we performed prognostic modeling for patients with early- and late-stage EEC, respectively. Importantly, we developed and validated an effective prognostic model for patients with early-stage EEC, which, to the best of our knowledge, is the first quantitative protein-marker-driven model for this patient population. Although the patients represented in the training and validation cohorts may have some intrinsic differences (e.g., a high rate of obesity among the patients in the United States and different age distributions between the cohorts), the risk group index defined by our model consistently shows an improved discriminating power across the two patient cohorts using different cutoff values, which highlights the robustness of our model. In contrast, major conventional prognostic factors (e.g., disease stage and tumor grade) appear to have little power

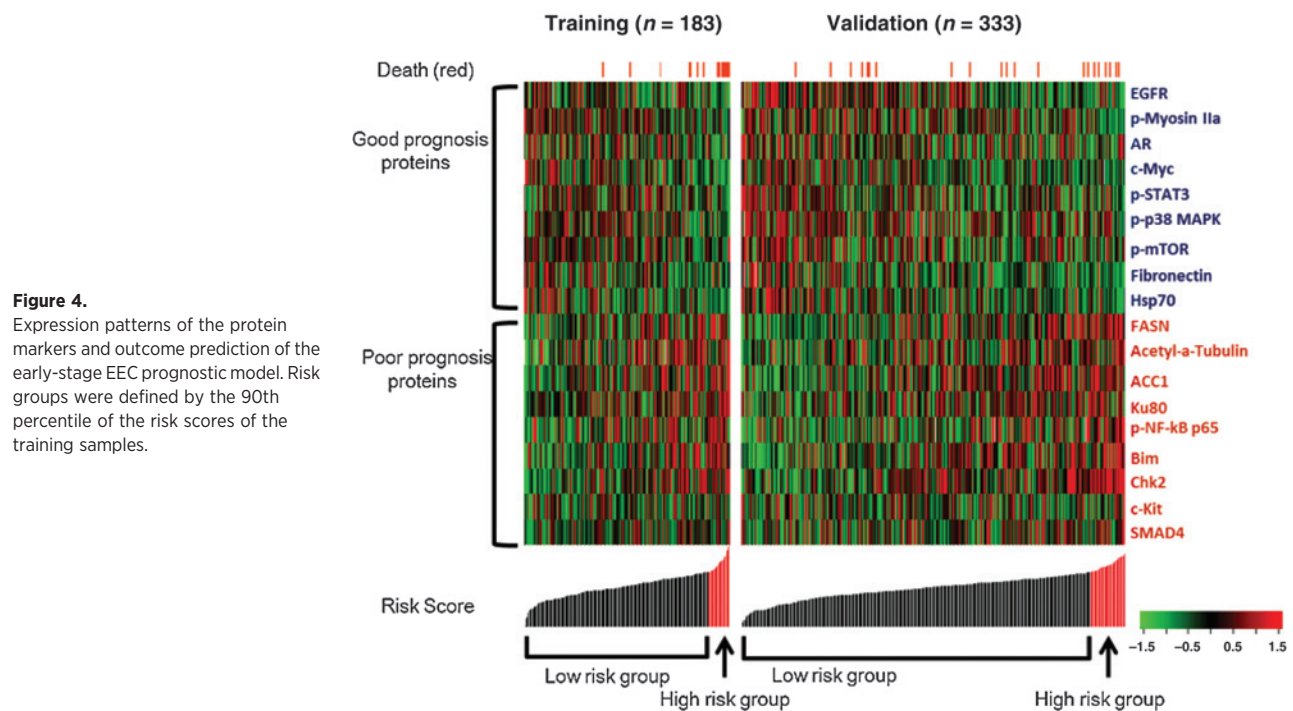


Figure 4. Expression patterns of the protein markers and outcome prediction of the early-stage EEC prognostic model. Risk groups were defined by the 90th percentile of the risk scores of the training samples.

for this patient population. Therefore, our integrative model for early-stage EEC is of potential clinical utility for identifying/prioritizing the patients with early-stage EEC who have a high risk of disease recurrence and death.

Our integrative model for early-stage EEC contains two clinical features and 18 protein markers. As expected, older age and higher tumor grade are associated with a worse prognosis. As shown in Fig. 4, there are two distinct groups of protein markers in regard to prognosis. The proteins for which the expression levels are associated with a good prognosis include EGFR, myosin IIa, AR, c-Myc, STAT3, p38 MAPK, mTOR, fibronectin, and HSP70. Among these markers, STAT3, MAPK, and mTOR are downstream targets of EGFR signaling. The prognostic effect of EGFR expression in cancer has been controversial in the literature; and it often depends on the specific tumor context, which may be due to the complexities of the downstream effects of EGFR signaling. Our model indicates that a high EGFR level in early-stage EEC is associated with a good prognosis. Indeed, a dual role of EGFR has been reported in endometrial cancer: high EGFR expression in well-differentiated EEC is associated with a low tumor grade and a favorable outcome; whereas high EGFR expression appears to promote disease progression in more aggressive, undifferentiated nonendometrioid tumors (35). In bladder cancer, a high level of EGFR expression has been reported to be a favorable prognostic factor (36); while in ovarian cancer, a high level of EGFR expression has been reported to be associated with a poor disease-free survival (37). A high level of expression of the androgen receptor has been reported to be a good prognostic marker for breast cancer and for all breast tumors and the triple-negative subtype (38, 39). This observation is potentially related to the tumor differentiation status. The protein markers for which high expression levels are associated with poor prognoses include FASN, acetyl- α -tubulin, ACC1, Ku80, NF- κ B, Bim, Chk2, c-Kit, and SMAD4. Among them, FASN and ACC1 are both related to fatty acid synthesis, and FASN

has been shown to be associated with poor patient survival in urothelial and colorectal cancers (40, 41). Ku80 and Chk2 are DNA damage sensors that may reflect the underlying mechanisms that drive tumors associated with poor outcomes; BIM is regulated by EGFR that is a marker of good prognosis in our model, and this observation may reflect decreased EGFR signaling in the patients with poor prognoses.

In terms of clinical applications, this model can be rapidly translated to the clinic using the RPPA platform, which has been made CLIA compliant in several centers and others are moving to make it CLIA compliant. With the relatively small number of 18 protein markers, it is also possible to develop a Luminex assay that would allow rapid transition to the clinic. Finally, a number of multiplex IHC approaches are now in place. Thus, the predictors can be readily transitioned to the clinic. Further evaluation using other independent cohorts is essential to establish the ultimate utility of our integrative model.

Patients diagnosed with late-stage EEC (stages III and IV) are usually prescribed aggressive treatment options such as chemotherapy. A potential utility of prognostic modeling for these patients is to identify the individuals with low-risk disease who may not need additional treatment. However, our integrative model did not provide additional prognostic power for this determination, which may be primarily due to the limited sample size. Another issue is the unbalanced event rates between training and validation sets. Larger patient cohorts will be needed to assess the clinical utility of the integrative model for late-stage EEC.

Our study has several limitations. First, we focused on overall survival as the primary endpoint in our analysis. Although overall survival is often used in the field (42), cancer-specific survival would be a more appropriate endpoint for such a purpose. Unfortunately, cancer-specific survival data are not available in our validation cohort. To minimize the limitation of overall survival (due to cancer-unrelated deaths), survival times were

Yang et al.

capped at 5 years as in the previous study (22). It would be of high clinical interest to further develop prognostic and predictive models for cancer-specific, disease free and metastasis-free survival based on our patient cohorts when such survival data are available. Second, our analysis only included a limited number of clinical variables due to incomplete clinical information of our patient samples. An important extension would be to integrate other types of clinical variables such as myometrial invasion and lymph node status. Third, we did not include treatment effects in our analysis. Adjuvant therapy (e.g., chemotherapy and/or radiation) may impact survival and thus confound the impact of prognostic factors or markers on survival. Unfortunately, treatment information in our patient cohorts is very incomplete, which prevented us from including these data in the analysis. However, in general, EEC patients with the same or similar stage/grade received similar (adjuvant) therapies. Thus, because (i) our modeling was carried out only for patients with similar disease stages (i.e., early-stage or late-stage) and (ii) the grade effect has been included in our model, this confounding effect would be relatively limited. In future, it is critical to validate our model based on the patient cohorts with high-quality and complete clinical information (such as those from clinical trials). Last but not the least, the current RPPA platform covers only about 200 proteins and therefore we may miss some important protein markers. We are in the process of extending our platform to many more proteins, which may further improve the prognostic power of the model. It remains unclear how several protein markers in our early-stage model, such as myosin IIa, fibronectin, and HSP70, are related to tumor progression of EEC and affect the clinical outcomes. Further efforts are required to elucidate the role of these proteins in the context of early-stage EEC. However, this mechanistic investigation can be pursued independent of the prognostic values conferred by these markers. Currently, only RPPA data are available for both training and validation cohorts, and we will include molecular features such as somatic mutations and copy number alterations, when the related data are available.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin* 2015;65:5–29.
2. Siegel R, Ward E, Brawley O, Jemal A. Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA Cancer J Clin* 2011;61:212–36.
3. Amant F, Moerman P, Neven P, Timmerman D, Van Limbergen E, Vergote I. Endometrial cancer. *Lancet* 2005;366:491–505.
4. Westin SN, Broaddus RR, Deng L, McCampbell A, Lu KH, Lacour RA, et al. Molecular clustering of endometrial carcinoma based on estrogen-induced gene expression. *Cancer Biol Ther* 2009;8:2126–35.
5. Salvesen HB, Haldorsen IS, Trovik J. Markers for individualised therapy in endometrial carcinoma. *Lancet Oncol* 2012;13:e353–61.
6. Gehrig PA, Bae-Jump VL. Promising novel therapies for the treatment of endometrial cancer. *Gynecol Oncol* 2010;116:187–94.
7. Fleming GF, Brunetto VL, Cella D, Look KY, Reid GC, Munkarah AR, et al. Phase III trial of doxorubicin plus cisplatin with or without paclitaxel plus filgrastim in advanced endometrial carcinoma: a Gynecologic Oncology Group Study. *J Clin Oncol* 2004;22:2159–66.
8. Jutzi L, Hoskins P, Lim P, Aquino-Parsons C, Tinker A, Kwon JS. The importance of adjuvant chemotherapy and pelvic radiotherapy in high-risk early stage endometrial carcinoma. *Gynecol Oncol* 2013;131:581–5.
9. Creasman WT, Odicino F, Maisonneuve P, Quinn MA, Beller U, Benedet JL, et al. Carcinoma of the corpus uteri. FIGO 26th Annual Report on the Results of Treatment in Gynecological Cancer. *Int J Gynaecol Obstet* 2006;95:S105–43.
10. Keys HM, Roberts JA, Brunetto VL, Zaino RJ, Spirtos NM, Bloss JD, et al. A phase III trial of surgery with or without adjunctive external pelvic radiation therapy in intermediate risk endometrial adenocarcinoma: a Gynecologic Oncology Group study. *Gynecol Oncol* 2004;92:744–51.
11. Akbani R, Ng PK, Werner HM, Shahmoradgoli M, Zhang F, Ju Z, et al. A pan-cancer proteomic perspective on The Cancer Genome Atlas. *Nat Commun* 2014;5:3887.
12. Li J, Lu Y, Akbani R, Ju Z, Roebuck PL, Liu W, et al. TCPA: a resource for cancer functional proteomics data. *Nat Methods* 2013;10:1046–7.
13. Spurrier B, Ramalingam S, Nishizuka S. Reverse-phase protein lysate microarrays for cell signaling analysis. *Nat Protoc* 2008;3:1796–808.
14. Sheehan KM, Calvert VS, Kay EW, Lu Y, Fishman D, Espina V, et al. Use of reverse phase protein microarrays and reference standard development for molecular network analysis of metastatic ovarian carcinoma. *Mol Cell Proteomics* 2005;4:346–55.
15. Hennessy BT, Lu Y, Gonzalez-Angulo AM, Carey MS, Myhre S, Ju Z, et al. A technical assessment of the utility of reverse phase protein arrays for the

Authors' Contributions

Conception and design: J.-Y. Yang, H.B. Salvesen, H. Liang
Development of methodology: J.-Y. Yang, G.B. Mills, H. Liang
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): H.M.J. Werner, J. Li, S.N. Westin, M.K. Halle, J. Trovik, H.B. Salvesen
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): J.-Y. Yang, H.M.J. Werner, J. Li, H.B. Salvesen, G.B. Mills, H. Liang
Writing, review, and/or revision of the manuscript: J.-Y. Yang, H.M.J. Werner, S.N. Westin, J. Trovik, H.B. Salvesen, G.B. Mills, H. Liang
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): J. Li, M.K. Halle, J. Trovik, H.B. Salvesen, G.B. Mills, H. Liang
Study supervision: H.B. Salvesen, H. Liang
Other (performed RPPA analysis data in the article): Y. Lu

Acknowledgments

The authors gratefully acknowledge contributions from the Endometrial Cancer Working Group of TCGA Research Network. The authors thank Kadri Madisoo and Britt Edvardsen (the University of Bergen) for technical support and LeeAnn Chastain (MDACC) for article editing.

Grant Support

This study was supported by the NIH through grant number CA143883 to G.B. Mills; CA088084 K12 Calabresi Scholar Award to S.N. Westin; CA175486 to H. Liang; CA098258 SPORE in Uterine Cancer to G.B. Mills, S.N. Westin, H. Liang and CCSG grant CA016672 to MDACC. Additional support was received from Kumoh National Institute of Technology to J.-Y. Yang; Helse Vest, Research Council of Norway, The Norwegian Cancer Society (Harald Andersens legat) and University of Bergen to H.B. Salvesen; the Lorraine Dell Program in Bioinformatics for Personalization of Cancer Medicine, one Lee Clark Fellow Award from the Jeanne F. Shelby Scholarship Fund and a grant from the Cancer Prevention and Research Institute of Texas (RP140462) to H. Liang.

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 January 13, 2015; revised July 21, 2015; accepted July 23, 2015; published OnlineFirst July 29, 2015.

- study of the functional proteome in non-microdissected human breast cancers. *Clin Proteomics* 2010;6:129–51.
16. Yang JY, Yoshihara K, Tanaka K, Hatae M, Masuzaki H, Itamochi H, et al. Predicting time to ovarian carcinoma recurrence using protein markers. *J Clin Invest* 2013;123:3740–50.
 17. Tibes R, Qiu Y, Lu Y, Hennessy B, Andreeff M, Mills GB, et al. Reverse phase protein array: validation of a novel proteomic technology and utility for analysis of primary leukemia specimens and hematopoietic stem cells. *Mol Cancer Ther* 2006;5:2512–21.
 18. Lu Y, Ling S, Byers L, Coombes K, Mills G, Akbani R. Using Reverse Phase Protein Array (RPPA) as a Pharmacodynamic Assay for Functional Proteomics, Biomarker Discovery, and Drug Development in Cancer. *Semin Oncol*. In Press.
 19. Yuan Y, Van Allen EM, Omberg L, Wagle N, Amin-Mansour A, Sokolov A, et al. Assessing the clinical utility of cancer genomic and proteomic data across tumor types. *Nat Biotechnol* 2014;32:644–52.
 20. Salvesen HB, Carter SL, Mannelqvist M, Dutt A, Getz G, Stefansson IM, et al. Integrated genomic profiling of endometrial carcinoma associates aggressive tumors with indicators of PI3 kinase activation. *Proc Natl Acad Sci U S A* 2009;106:4834–9.
 21. The Cancer Genome Research Network. Integrated genomic characterization of endometrial carcinoma. *Nature* 2013;497:67–73.
 22. The Cancer Genome Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* 2011;474:609–15.
 23. Neeley ES, Komblau SM, Coombes KR, Baggerly KA. Variable slope normalization of reverse phase protein arrays. *Bioinformatics* 2009;25:1384–9.
 24. Yang JY, He X. A multistep protein lysate array quantification method and its statistical properties. *Biometrics* 2011;67:1197–205.
 25. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc B* 2005;67:301–20.
 26. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat softw* 2010;33:1.
 27. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization paths for Cox's proportional hazards model via coordinate descent. *J Stat Softw* 2011;39:1–13.
 28. Harrell F, Lee KL, Mark DB. Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361–87.
 29. Nicholl J, Jacques RM, Campbell MJ. Direct risk standardisation: a new method for comparing casemix adjusted event rates using complex models. *BMC Med Res Methodol* 2013;13:133.
 30. Klein JP, Andersen PK. Regression modeling of competing risks data based on pseudovalues of the cumulative incidence function. *Biometrics* 2005;61:223–9.
 31. Arnold D, Cook D. The intensive care unit. *Eur Respiratory Soc Monograph* 2013;62:85–9.
 32. Creutzberg CL, van Putten WL, Koper P, Lybeert ML, Jobsen JJ, Wárlám-Rodenhuis CC, et al. Surgery and postoperative radiotherapy versus surgery alone for patients with stage-1 endometrial carcinoma: multicentre randomised trial. *Lancet* 2000;355:1404–11.
 33. Wik E, Raeder MB, Krakstad C, Trovik J, Birkeland E, Hoivik EA, et al. Lack of estrogen receptor-alpha is associated with epithelial-mesenchymal transition and PI3K alterations in endometrial carcinoma. *Clin Cancer Res* 2013;19:1094–105.
 34. Schouten HJ. Adaptive biased urn randomization in small strata when blinding is impossible. *Biometrics* 1995;51:1529–35.
 35. Konecny G, Santos L, Winterhoff B, Hatmal M, Keeney G, Mariani A, et al. HER2 gene amplification and EGFR expression in a large cohort of surgically staged patients with nonendometrioid (type II) endometrial cancer. *Br J Cancer* 2008;100:89–95.
 36. Chakravarti A, Winter K, Wu C-L, Kaufman D, Hammond E, Parliament M, et al. Expression of the epidermal growth factor receptor and Her-2 are predictors of favorable outcome and reduced complete response rates, respectively, in patients with muscle-invasive bladder cancers treated by concurrent radiation and cisplatin-based chemotherapy: a report from the Radiation Therapy Oncology Group. *Int J Radiat Oncol Biol Phys* 2005;62:309–17.
 37. Skrimisdottir I, Seidal T, Sorbe B. A new prognostic model comprising p53, EGFR, and tumor grade in early stage epithelial ovarian carcinoma and avoiding the problem of inaccurate surgical staging. *Int J Gynecol Cancer* 2004;14:259–70.
 38. Yu KD, Zhu R, Zhan M, Rodriguez AA, Yang W, Wong S, et al. Identification of prognosis-relevant subgroups in patients with chemoresistant triple-negative breast cancer. *Clin Cancer Res* 2013;19:2723–33.
 39. Gonzalez-Angulo AM, Stemke-Hale K, Palla SL, Carey M, Agarwal R, Meric-Bertram F, et al. Androgen receptor levels and association with PIK3CA mutations and prognosis in breast cancer. *Clin Cancer Res* 2009;15:2472–8.
 40. Long QQ, Yi YX, Qiu J, Xu CJ, Huang PL. Fatty acid synthase (FASN) levels in serum of colorectal cancer patients: correlation with clinical outcomes. *Tumour Biol* 2014;35:3855–9.
 41. Hamada S, Horiguchi A, Asano T, Kuroda K, Asakuma J, Ito K, et al. Prognostic impact of fatty acid synthase expression in upper urinary tract urothelial carcinoma. *Jpn J Clin Oncol* 2014;44:486–92.
 42. Onsrud M, Cvancarova M, Hellebust TP, Trope CG, Kristensen GB, Lindemann K. Long-term outcomes after pelvic radiation for early-stage endometrial cancer. *J Clin Oncol* 2013;31:3951–6.

Clinical Cancer Research

Integrative Protein-Based Prognostic Model for Early-Stage Endometrioid Endometrial Cancer

Ji-Yeon Yang, Henrica M.J. Werner, Jie Li, et al.

Clin Cancer Res 2016;22:513-523. Published OnlineFirst July 29, 2015.

Updated version Access the most recent version of this article at:
[doi:10.1158/1078-0432.CCR-15-0104](https://doi.org/10.1158/1078-0432.CCR-15-0104)

Supplementary Material Access the most recent supplemental material at:
<http://clincancerres.aacrjournals.org/content/suppl/2015/08/08/1078-0432.CCR-15-0104.DC1.html>

Cited articles This article cites 41 articles, 10 of which you can access for free at:
<http://clincancerres.aacrjournals.org/content/22/2/513.full.html#ref-list-1>

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, contact the AACR Publications Department at permissions@aacr.org.