

Original Investigation

Identification and Construction of Combinatory Cancer Hallmark–Based Gene Signature Sets to Predict Recurrence and Chemotherapy Benefit in Stage II Colorectal Cancer

Shanwu Gao, MSc; Chabane Tibiche, MSc; Jinfeng Zou, PhD; Naif Zaman, MSc; Mark Trifiro, MD; Maureen O'Connor-McCourt, PhD; Edwin Wang, PhD

IMPORTANCE Decisions regarding adjuvant therapy in patients with stage II colorectal cancer (CRC) have been among the most challenging and controversial in oncology over the past 20 years.

OBJECTIVE To develop robust combinatory cancer hallmark–based gene signature sets (CSS sets) that more accurately predict prognosis and identify a subset of patients with stage II CRC who could gain survival benefits from adjuvant chemotherapy.

DESIGN, SETTING, AND PARTICIPANTS Thirteen retrospective studies of patients with stage II CRC who had clinical follow-up and adjuvant chemotherapy were analyzed. Respective totals of 162 and 843 patients from 2 and 11 independent cohorts were used as the discovery and validation cohorts, respectively. A total of 1005 patients with stage II CRC were included in the 13 cohorts. Among them, 84 of 416 patients in 3 independent cohorts received fluorouracil-based adjuvant chemotherapy.

MAIN OUTCOMES AND MEASURES Identification of CSS sets to predict relapse-free survival and identify a subset of patients with stage II CRC who could gain substantial survival benefits from fluorouracil-based adjuvant chemotherapy.

RESULTS Eight cancer hallmark–based gene signatures (30 genes each) were identified and used to construct CSS sets for determining prognosis. The CSS sets were validated in 11 independent cohorts of 767 patients with stage II CRC who did not receive adjuvant chemotherapy. The CSS sets accurately stratified patients into low-, intermediate-, and high-risk groups. Five-year relapse-free survival rates were 94%, 78%, and 45%, respectively, representing 60%, 28%, and 12% of patients with stage II disease. The 416 patients with CSS set–defined high-risk stage II CRC who received fluorouracil-based adjuvant chemotherapy showed a substantial gain in survival benefits from the treatment (ie, recurrence reduced by 30%-40% in 5 years).

CONCLUSIONS AND RELEVANCE The CSS sets substantially outperformed other prognostic predictors of stage 2 CRC. They are more accurate and robust for prognostic predictions and facilitate the identification of patients with stage II disease who could gain survival benefit from fluorouracil-based adjuvant chemotherapy.

JAMA Oncol. doi:10.1001/jamaoncol.2015.3413
Published online October 22, 2015.

← Editorial

+ Supplemental content at
jamaoncol.com

Author Affiliations: Honghe University, Yunnan Province, China (Gao); National Research Council Canada, Montreal, Quebec, Canada (Gao, Tibiche, Zou, Zaman, O'Connor-McCourt, Wang); Department of Medicine, McGill University, Montreal, Quebec, Canada (Trifiro, Wang); Jewish General Hospital, Montreal, Quebec, Canada (Trifiro); McGill University Center for Bioinformatics, Montreal, Quebec, Canada (Wang).

Corresponding Author: Edwin Wang, PhD, Department of Medicine, McGill University, 1001 Decarie Blvd, Montreal, QC, Canada, H4A 3J1 (edwin.wang@mcgill.ca)

Decisions regarding adjuvant therapy (ie, chemotherapy given after surgery to reduce the risk of cancer recurrence) in patients with stage II colorectal cancer (CRC) have been among the most challenging¹⁻⁵ and controversial in oncology over the past 20 years. Moreover, professional clinical organizations disagree about the clinical guidelines for stage II disease. The Scottish Intercollegiate Guidelines Network does not favor adjuvant therapy⁶; the American Society of Clinical Oncology suggests that adjuvant chemotherapy should be considered⁷; and the National Cooperative Cancer Network recommends that adjuvant therapy can be applied for patients with stage II disease and high-risk features.⁸ However, there have thus far been no solid clinicopathologic features or biomarkers for identifying high-risk patients with stage II disease who could benefit from adjuvant therapy. As a result, patients and physicians are often uncomfortable forgoing adjuvant therapy.¹

Despite numerous clinical trials and meta-analyses, there is no solid evidence showing that chemotherapy given after surgery for stage II CRC improves survival; therefore, the benefit of adjuvant chemotherapy for patients with stage II disease remains a matter of debate.^{1,9,10} In an effort to clarify the benefit associated with adjuvant therapy in stage II disease, O'Connor et al¹¹ analyzed data from nearly 25 000 patients with stage II disease in the Surveillance, Epidemiology, End Results and Medicare registry database and concluded that there was no difference in 5-year survival between those who received postoperative chemotherapy and those who did not. They further showed that clinicopathologic features alone are not a sufficient basis for patient treatment selection; this result has been demonstrated in many other studies of patients with stage II cancer.¹⁻⁵

Over the past decade, extensive efforts have been made to identify prognostic gene signatures for patients with stage II disease using gene expression profiles. However, most of the gene signatures are not robust (ie, a gene signature loses its predictive power for a given independent patient cohort).^{12,13} For example, Park et al¹³ tested 5 popular genomic predictors for CRC prognosis using 2 independent cohorts and found that only 2 of them showed robust performance. Although the predictions of the Oncotype DX Colon Cancer test (hereafter “Oncotype DX”)¹⁴ and ColoPrint (Agendia BV)^{15,16} are reproducible, they are not predictive of adjuvant treatment benefits for patients with stage II disease. This conclusion is further supported by independent validation studies of several thousand patients with stage II disease.¹⁴⁻¹⁹ So far, none of the gene signatures has been able to predict which patients with stage II disease could benefit from adjuvant therapy or to assist in guiding treatment decisions, a primary goal of personalized medicine in oncology.

In the present study, we modified our research group's previously developed Multiple Survival Screening (MSS) algorithm,¹² which is able to identify cancer hallmark-based gene signatures, by constructing combinatory cancer hallmark based gene expression signature sets (CSS sets) for accurately predicting the prognosis of patients with stage II disease. We showed that CSS sets successfully predicted the recurrence and adjuvant therapeutic benefits in patients with

At a Glance

- No biomarkers have been able to accurately predict which patients with stage II colorectal cancer (CRC) would gain benefits from adjuvant therapy and thus guide treatment decisions.
- We identified and validated combinatory cancer hallmark-based gene signature sets (CSS sets) for accurately determining prognosis and adjuvant chemotherapy benefits in about 1000 patients with stage II CRC from 13 independent cohorts.
- The CSS sets were used to stratify patients with stage II CRC into low-, intermediate-, and high-risk groups with 5-year relapse-free survival rates of 94%, 78%, and 45%, respectively.
- The CSS set-defined patients with high-risk stage II CRC gained significant survival benefits from fluorouracil adjuvant chemotherapy (ie, reduced recurrence by 30%-40% in 5 years; $P = .004$).
- The CSS set-defined patients with low-risk and intermediate-risk stage II CRC did not gain survival benefits, but instead experienced shorter survival after fluorouracil adjuvant chemotherapy ($P = .04$ and $P = .003$, respectively).

stage II disease. To our knowledge, this is the first study to show that patients with gene expression signature-defined high-risk stage II CRC could achieve significant survival benefits through adjuvant therapy. These results shed light on the controversial issue for stage II CRC treatment, which has been debated for more than 20 years. The CSS sets could be useful for making treatment decisions for patients with stage II disease.

Methods

Patients and Tumor Samples

More than 1000 clinically annotated stage II CRC tumor samples from 13 independent cohorts²⁰⁻³⁰ were used for our analysis: Gene Expression Omnibus (GEO) microarray data repository Nos. GSE37892, GSE17538, GSE14333, GSE33113, GSE39582, GSE21510, GSE26906, GSE27854, GSE12945, GSE41258, GSE16125, GSE24551, and GSE12032 (<http://www.ncbi.nlm.nih.gov/geo/>). These samples were discovered by extensively searching microarray databases and were chosen for analysis based on the availability of clinically annotated data (minimum inclusion criteria, information on either recurrence/metastasis or overall survival). Clinical and pathologic data and molecular features were extracted from the GEO data sets and associated publications.²⁰⁻³⁰ All patients who provided the tumor samples were monitored for either relapse (distant metastases or locoregional recurrence) or overall survival (median follow-up times, 63.5 months for the training set and 53.0 months for the validation set) (see eTables 1 through 16 in the Supplement).

Among the GEO groups, GSE14333, GSE39582, and GSE17538 contained the samples of patients who had received adjuvant chemotherapy. Adjuvant chemotherapy information for GSE17538 has been updated in GSE29623.³¹ Fluorouracil-based adjuvant chemotherapy (fluorouracil as a key component, single-agent fluorouracil, fluorouracil and oxaliplatin, fluorouracil and folic acid, or fluorouracil and others) had been administered to the patients of 3 independent

cohorts: GSE14333, GSE37892 and GSE17538. Four microarray platforms were used by these cohorts (Affymetrix HG-U133, Affymetrix HG-133A, Affymetrix Human Exon 1.0, and Hitachisoft AceGene Human Oligo Chip). Data processing and normalization techniques are detailed in eMethods in the [Supplement](#).

Samples (n = 162) from 2 independent training cohorts (GSE37892 and GSE33113) were used for identifying cancer hallmark-based gene signatures and constructing CSS sets. Stage II CRC samples (n = 767) of the patients who did not receive adjuvant chemotherapy in 11 independent cohorts were used for the validation of the CSS sets. Stage II CRC samples (n = 416) of the patients who received adjuvant chemotherapy in 3 independent cohorts were used to examine the survival benefits of fluorouracil-based adjuvant chemotherapy in the high-risk group identified by the CSS sets.

Generating Cancer Hallmark-Based Gene Signatures

To generate the cancer hallmark-based gene signatures from the training set, we followed the procedures of our group's previously developed MSS algorithm¹² using the stage II CRC samples from 2 independent training cohorts (GSE37892 and GSE33113). The pseudocode of the algorithm is provided in eMethods in the [Supplement](#).

Determining Prognosis From New Samples or Drug-Treated Samples Using the CSS Sets

The procedure for construction of the CSS sets using cancer hallmark-based gene signatures is provided in eMethods in the [Supplement](#). The prediction procedure is similar to the leave-1-out cross-validation procedure (eMethods in the [Supplement](#)) except that we only used the untreated samples to construct centroids for gene signatures. Briefly, the nearest-shrunken-centroid method was used to calculate "average" feature vectors, V_{low} and V_{high} for the low-risk and high-risk tumor samples, respectively, from each gene signature of the samples (stage I and II together) that had not been treated with drugs. For a drug-treated stage II sample, we extracted its feature vectors from each gene signature using the sample's gene expression profile. Using the Pearson correlation coefficients between the sample's feature vectors and V_{low} and V_{high} as well as the signature-predicting rules used in the leave-1-out cross-validation procedure, we then assigned the samples to high-, intermediate-, or low-risk groups.

For a given new stage II sample that had been preferably profiled using Affymetrix arrays, we meta-normalized it with a reference data set (eg, GSE17538, GSE37892, GSE14333). The same prediction procedure would be applied to predict the new sample.

Statistical Analysis

Statistical significance of the prognostic groups (ie, high-, intermediate-, or low-risk groups defined by CSS sets) was determined using Kaplan-Meier survival plots. A prognostically significant result was defined by log-rank $P < .05$. Prognostic significance of clinicopathologic factors and molecular features (ie, mutation status of *BRAF*, *APC*, *TP53*, and *KRAS* or mismatch repair status) were performed with the use of the Cox

proportional hazards regression model. P values were based on likelihood ratio tests. All the analyses were performed using the statistical R packages.

Results

Inadequacy of Single Gene Signatures

In this study, we extensively collected 13 public microarray data sets containing more than 1000 stage II CRC samples containing clinical follow-up information. To identify predictive gene signatures for patients with stage II disease, we ran the MSS algorithm,¹² designed to identify robust cancer biomarkers by focusing on cancer hallmark-associated genes. In choosing a training set, we had several considerations. (1) The clinical information should include time to recurrence. (2) Among the 13 collected data sets, 8 sets used the Affymetrix HG-U133 array platform, and so to facilitate validation studies, we preferred to take training sets that used this array platform. (3) To test gene signatures for predicting adjuvant treatment benefits, we excluded the data sets that contained drug-treated samples as training sets.

Adhering to these criteria, we randomly selected GSE37892 as a training set. We took stage II samples (n = 73) from GSE37892 and ran MSS to first perform a survival test for each gene and then group the survival-significant genes ($P < .05$) based on cancer hallmark-associated Gene Ontology (GO) terms (eg, cell cycle, apoptosis; see eTable 17 in the [Supplement](#)).^{12,32-35} The procedure is detailed in eMethods in the [Supplement](#). Briefly, for a cancer hallmark GO-defined gene group, we focused on 60 to 100 modulated genes between recurred and nonrecurred samples. From the training set, we generated 36 random data sets by randomly picking up 70% of the original training samples.

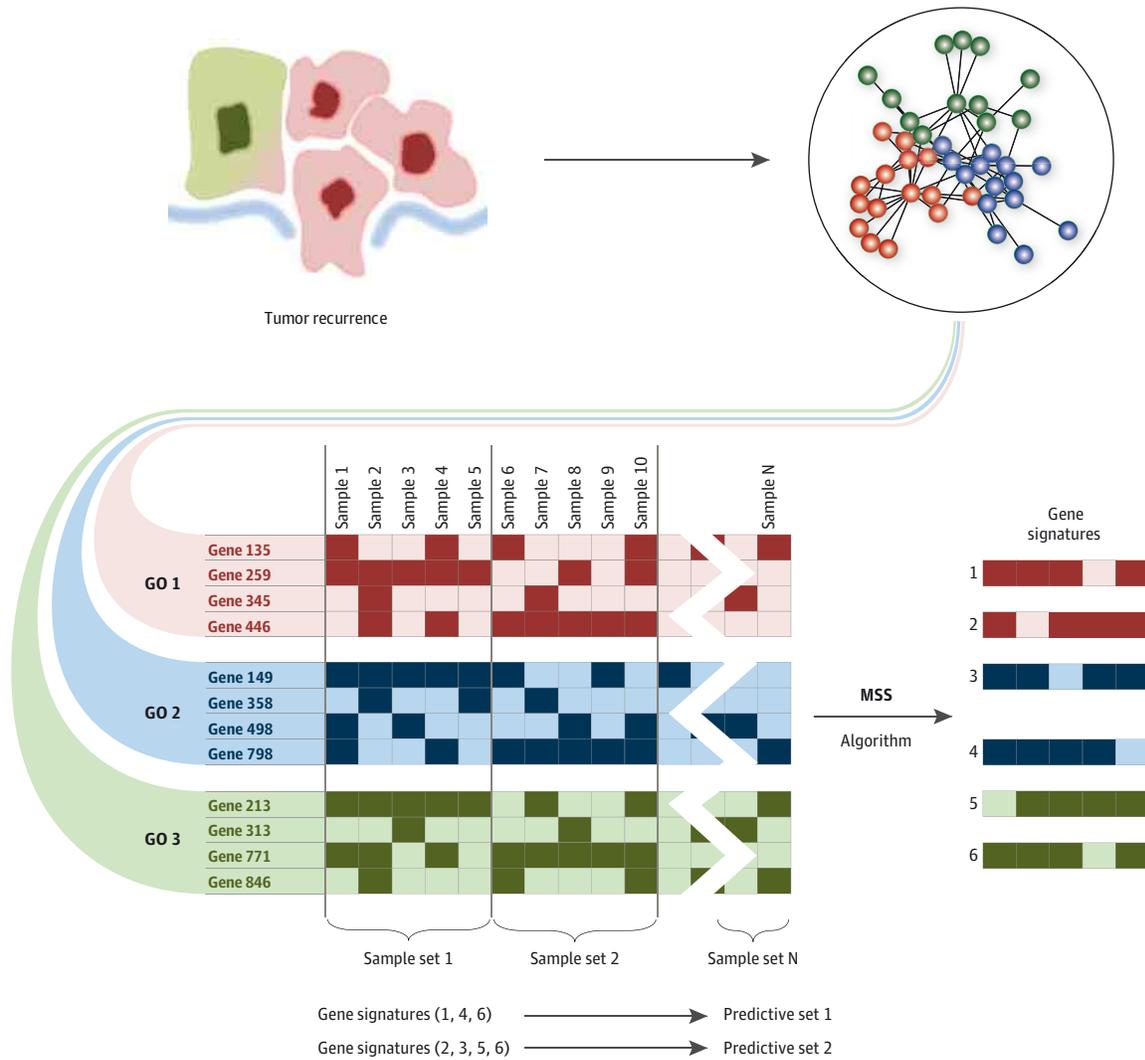
Meanwhile, we generated 1 million random gene sets (30 genes per set) by randomly picking from the GO-defined genes. From 1 million random gene sets, we collected 1000 to 5000 random gene sets that could distinguish low- from high-risk groups ($P < .05$) across more than 90% of the 36 random data sets for a cancer hallmark GO-defined-gene group ($P < .005$). We then used MSS to identify a set of genes as a gene signature (eMethods in the [Supplement](#)). To identify more gene signatures, we extended these procedures to run all of the samples (73 and 57 samples are stage II and III, respectively) from GSE37892.

In total, 8 cancer hallmark gene signatures were identified (eTable 18 in the [Supplement](#)). We used leave-1-out cross-validation (eMethods in the [Supplement](#)) to test these gene signatures in 12 other independent patient cohorts of patients with stage II disease and found that they predicted prognosis but failed to predict adjuvant treatment benefits. These results prompted us to think about a new strategy of cancer biomarker discovery.

Improved Prediction of Adjuvant Treatment Effect Using CSS Sets

Cancer traits (eg, cancer recurrence, metastasis) are complex on least at 2 levels: (1) for a given sample, several biological pro-

Figure 1. Overview of the Combinatory Cancer Hallmark-Based Gene Signature Set Approach



Tumor recurrence (ie, a cancer hallmark trait) can be represented by a metastasis-based molecular network.³⁴ In the network, several biological processes (eg, nodes marked with different colors, each color representing a biological process). Each biological process can be represented by a Gene Ontology (GO) term, and each GO term has a set of genes associated, but not all of the genes associated with a GO term are activated in a single tumor sample. The heatmap (checkerboarded area) shows that genes are activated in different

tumor samples. Therefore, a gene signature, which is identified from a hallmark GO term using the Multiple Survival Screening (MSS) algorithm,¹² can only represent a fraction of the tumor samples. Multiple combinations of the gene signatures will finally predict the prognoses of most of the tumor samples (eg, sample sets 1, 2,... N). Nodes and links in the network represent genes and their interactions. The dark colors in the heatmap represent genes that are highly activated and used by tumors.

cesses are involved in the associated trait; and (2) for different samples of even a same cancer type or subtype, genes used from the same biological process could be different (Figure 1). Several research teams have explored and elucidated these complexities,^{34,36-38} and from these insights, we constructed CSS sets, each of them containing several distinct molecular mechanism-based (ie, MSS-derived cancer hallmark) gene signatures to boost the prediction performance. For predicting tumor recurrence, a cancer hallmark gene signature represents a biological process that is part of the molecular mechanism of tumor recurrence. Thus, collaborative cancer hallmark gene signatures within a CSS set could foster greater cohesion and interactions and thus could increase both pre-

dition accuracy and recall rate by unifying the predictions from multiple CSS sets (Figure 1). Procedures for constructing CSS sets are detailed in eMethods in the Supplement. Similarly, cancer-related biological pathways could be used to build CSS sets.

We built CSS sets by examining the stage II CRC samples pooled from both GSE37892 (the training set) and GSE33133 (independent set) to avoid bias toward the training set. We found that for predicting low-risk samples, the best results were obtained when any 4 of the 8 MSS-derived gene signatures had consensus predictions. For the high-risk samples, the best results were obtained when all 8 MSS-derived gene signatures had consensus predictions (eTable 19 in the Supplement). We

Table. CSS Set Prediction Accuracies and Recall Rates for Patients With Stage II Colorectal Cancer^a

GEO Cohort Data Set ^b	Samples, No.	Risk, %					
		Low		Intermediate		High	
		Accuracy ^c	Recall ^d	Accuracy ^c	Recall ^d	Accuracy ^e	Recall ^f
GSE37892 (training)	73	96.2	78.5	90.9	15.4	55.6	62.5
GSE33113 (training)	90	94.7	50.7	72.7	33.8	38.9	38.9
GSE26906	90	90.7	71.0	86.7	18.8	57.1	38.1
GSE17538	72	100.0	66.0	94.4	32.1	88.9	88.9
GSE39582	264	89.5	74.8	78.6	20.8	69.6	41.0
GSE21510	54	100.0	93.5	33.3	6.5	100.0	33.3
GSE14333	94	97.7	64.6	86.4	29.2	42.9	42.9
GSE27854	41	100.0	79.4	77.8	20.6	100.0	71.4
GSE12945 and GSE41258	72	93.8	67.2	87.5	10.4	NA	NA
GSE16125 and GSE24551	95	84.3	64.2	64.7	32.8	80.0	28.6
GSE12032	92	88.1	95.2	13.6	4.5	100.0	10.0

Abbreviations: CSS set, combinatory cancer hallmark-based gene signature set; GEO, Gene Expression Omnibus microarray data repository; NA, no applicable (when <6 recurred samples were present in a data set, the accuracy and recall rates of the high-risk group was not calculated).

^a Recall is the fraction of labeled instances (ie, true low- or high-risk samples) that are retrieved.

^b When a data set contained fewer than 30 samples, it was meta-normalized with another data set, and both data sets shared the same microarray platform. If a sample had no recurrence/metastasis information, we excluded it from our analysis.

^c Percentage of nonrecurred (ie, true low-risk) samples in the predicted low-risk group.

^d Percentage of the predicted low-risk samples from the nonrecurred group.

^e Percentage of recurred (ie, true high-risk) samples in the predicted high-risk group.

^f Percentage of the predicted high-risk samples from the recurred group.

further tested the CSS sets in 767 stage II CRC samples of 11 independent cohorts where drug-treated samples were removed using the leave-1-out cross-validation approach (eMethods in the Supplement). The CSS sets assigned 60%, 28%, and 12%, respectively, of all the stage II disease into low-, intermediate-, and high-risk groups, with 5-year relapse-free survival rates of 94%, 78%, and 45%, respectively ($P = .02$ to $P < .001$).

As detailed in the Table, Figure 2, and eFigure 1 in the Supplement, we predicted low-risk patients with 94% accuracy (median prediction accuracy of the 11 independent validation cohorts, defining accuracy as the percentage of actual low-risk patients found to be in the CSS set-defined low-risk group). We also significantly increased the prediction accuracy for high-risk samples to 55%.

Of note, low-risk prediction accuracies were similar among the validation cohorts, whereas high-risk prediction accuracies varied from data set to data set. These results could be owing to the clinical variability of the tumors (eTables 1-16 in the Supplement) or the sample size differences of the recurred and nonrecurred samples between the data sets. If a group (ie, recurred or nonrecurred group) in a data set had a small sample size, 1 or 2 uncorrected predictions would dramatically affect the prediction accuracy of that group. Indeed, the sample sizes were very different between nonrecurred (30-60 samples) and recurred (10-20 samples) stage II CRC samples in most data sets (eTable 20 in the Supplement), which explains the variable high-risk prediction accuracy between data sets.

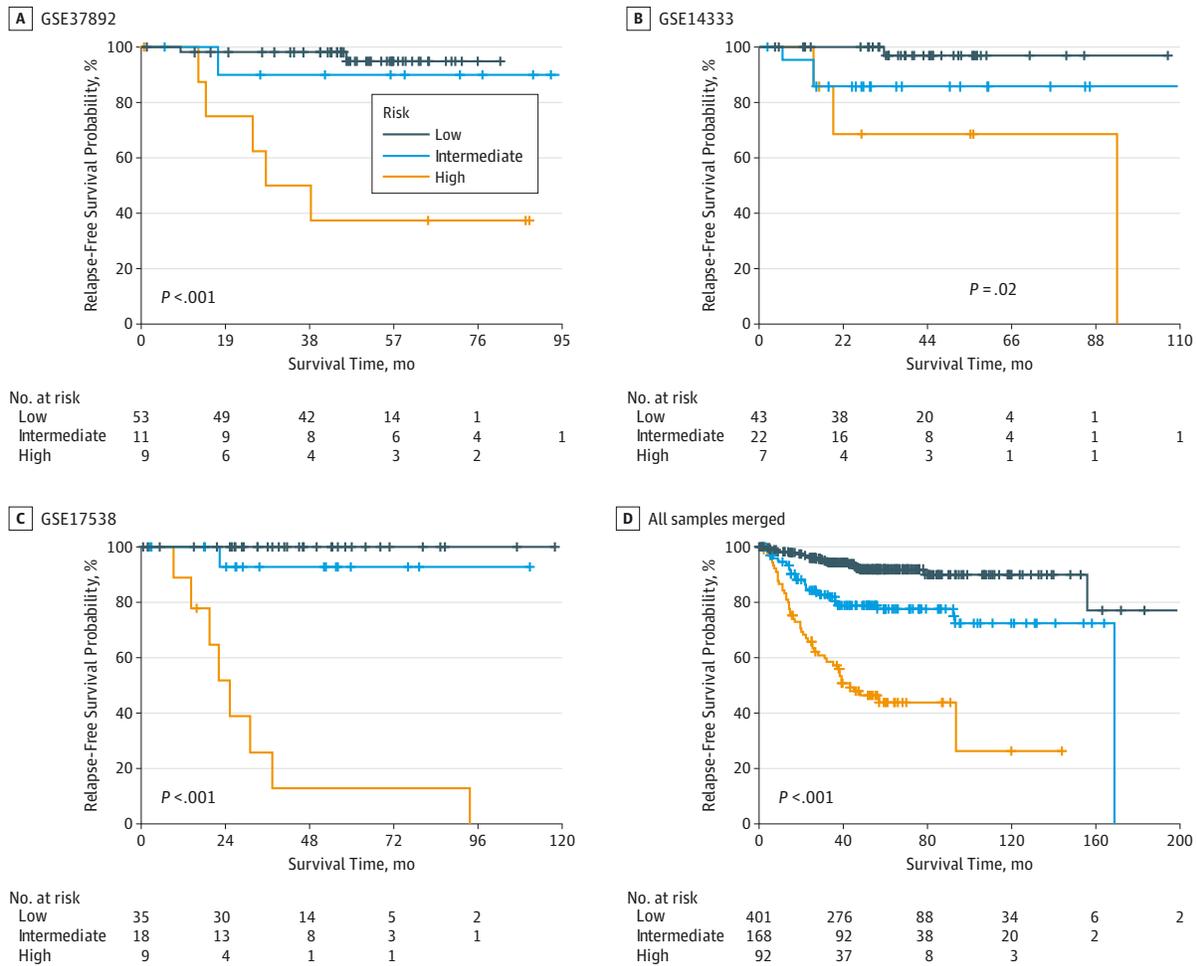
In addition, the microarray platform used in the cohort also affected accuracy. Low-risk prediction accuracy was 96% in Affymetrix HG-U133/Affymetrix HG-133A cohorts, 84% for Af-

fymetrix Human Exon 1.0, and 88% for Hitachisoft AceGene Human Oligo Chip. Because the training sets used the Affymetrix HG-U133 arrays, these results suggest that the prediction performance of the CSS sets tends to be better when the validation sets use the same array as the training sets.

The 94% low-risk prediction accuracy from the CSS sets is significantly higher than those from Oncotype DX (87%)¹⁴ and ColoPrint (88%).^{15,16} Because low-risk patients do not need to receive adjuvant treatment, it is critical to have high prediction accuracy for low-risk patients to make biomarkers clinically useful. Remarkably, the 55% prediction accuracy of the CSS sets for the high-risk group is 2 to 3 times higher than those of Oncotype DX (22%)¹⁴ and ColoPrint (22%-26%).^{15,16} The highly enriched recurred samples in the CSS set-defined high-risk group provide an opportunity for examining the adjuvant therapy benefit for high-risk samples of stage II CRC.

To compare the prediction performance of the CSS sets with that of clinical factors and molecular features, we conducted relapse-free survival analysis of clinical factors and molecular features such as mutation status of important genes, mismatch repair status, and other molecular features using the Cox proportional hazards regression model (eTable 21 in the Supplement). We extended this analysis to the CSS set-defined low- and high-risk groups of all 6 independent validation cohorts in which patients had follow-up time. Among clinical molecular features, only pT4 staging predicted poor prognosis in stage II CRC (hazard ratio [HR], 2.6; 95% CI, 1.4-4.7; $P = .002$). It is no surprise that the tumors with the most advanced histologic category for local invasion, pT4, have a poor prognosis in stage II CRC, which has also been reported by others.³⁹ Clearly we demonstrated that the CSS sets have

Figure 2. Disease-Free 5-Year Kaplan-Meier Survival Curves of the 3 CSS Set-Defined Risk Groups for Patients With Stage II Colorectal Cancer



CSS Set indicates combinatory cancer hallmark-based gene signature set. The measured tumor samples came from 3 independent cohorts (GSE37892, GSE14333, and GSE17538) and a combination thereof from the Gene Expression

Omnibus microarray data repository. Drug-treated samples were excluded from the analysis as were cohorts that contained only metastasis information but no follow-up time. P values were obtained from the χ^2 test.

much better prediction performance (HR, 6.5; 95% CI, 4.1-10.3; $P < .001$) than pT4 staging. Moreover, histologic parameters to identify the features of pT4 are not always entirely straightforward, making recognition of pT4 stage difficult at times.^{39,40}

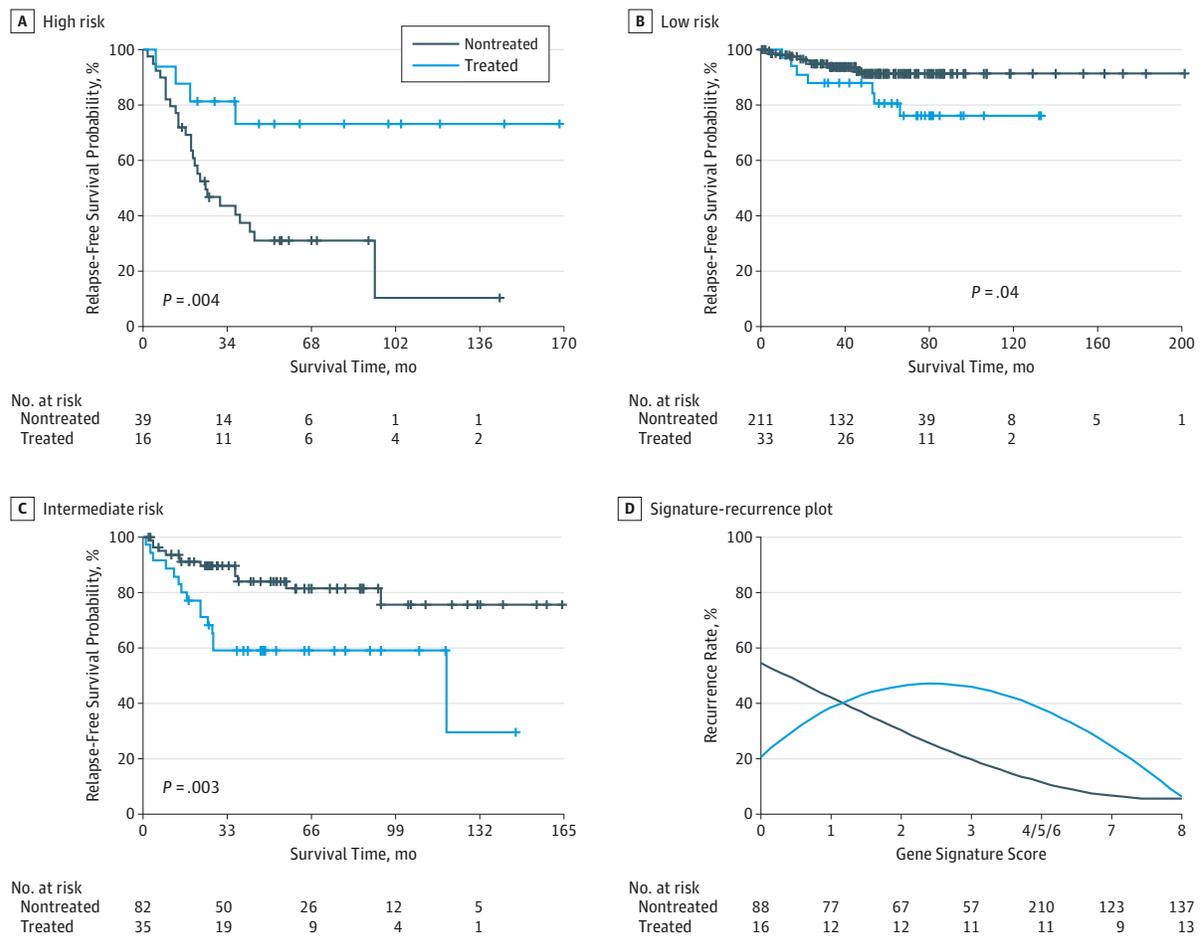
Identifying Patients With Stage II CRC Who Will Benefit From Fluorouracil-Based Adjuvant Chemotherapy

To evaluate whether the CSS sets are useful for guiding adjuvant therapy for patients with stage II disease, we analyzed samples from 3 cohorts (GSE14333, GSE17538, and GSE39582), each of which contained samples with and without fluorouracil-based adjuvant chemotherapy. The CSS sets were used to predict low-, intermediate- and high-risk patients. A comparison analysis showed that for the predicted high-risk group, fluorouracil-treated patients had markedly improved outcomes compared with patients who did not receive the treatment (Figure 3A; $P = .004$). Fluorouracil treatment reduced the recurrence by 30% to 40% in 5

years (Figure 3A). When examining more cohorts of fluorouracil adjuvant chemotherapy together, we found that the survival benefit gain of the treated patients was significantly increased (Figure 3A and eFigure 2A and B in the Supplement) ($P = .04$, $P = .04$, and $P = .004$, respectively, for the samples of cohorts GSE39582, GSE39582 + GSE14333, and GSE39582 + GSE14333 + GSE17538). To our knowledge, this is the first study to demonstrate that a fraction of patients with stage II disease could gain significant survival benefits from chemotherapy.

When extending the same analysis to the predicted low-risk groups, we found that patients with stage II disease did not gain significant survival benefit from fluorouracil-based adjuvant chemotherapy (eFigure 2C and D in the Supplement). When examining the samples from all 3 cohorts together, we found fluorouracil-treated low-risk patients to have significantly shorter survival than nontreated ones (Figure 3B) ($P = .04$). Notably, for the CSS set-predicted intermediate-risk groups, fluorouracil-treated patients with stage II dis-

Figure 3. Kaplan-Meier Plots and Linear Fit Plot of the 3 CSS Set-Defined Risk Groups for Distant Recurrences in Patients With Stage II Colorectal Cancer Treated With Fluorouracil or Nontreated.



CSS Set indicates combinatory cancer hallmark-based gene signature set. The measured tumor samples came from independent cohorts combinations thereof from the Gene Expression Omnibus (GEO) microarray data repository. A, High-risk patients from GEO GSE39582, GSE14333, and GSE17538. B, Low-risk patients from GEO GSE39582, GSE14333, and GSE17538.

C, Intermediate-risk patients from GEO GSE39582, GSE14333, and GSE17538. D, Linear fit of the likelihood of recurrence as a continuous function of gene signature score for the treated and nontreated samples; the treated samples that had scores of 4, 5, and 6 were merged into a single data point because their sample sizes were too small. P values were obtained from the χ^2 test.

ease had significantly shorter survival than nontreated ones (Figure 3C and eFigure 2E and F in the Supplement) ($P = .003$). These results suggest that these low-risk patients could have been spared the potentially toxic and costly effects of these treatments.

To explore the degree of benefit from chemotherapy in relation to the gene signatures as a continuous function, we assigned a gene signature score (GSS) to each sample in the validation sets based on the number of gene signatures that predicted the sample to be from a low-risk patient (eMethods in the Supplement). The likelihood of recurrence was fit as a linear function of the GSS for both fluorouracil-treated ($n = 84$) and untreated ($n = 767$) samples (Figure 3D). The higher GSS a sample had, the greater the possibility it was from a low-risk patient (eFigure 2G in the Supplement). When the GSS was between 0 to 1 (ie, any 7 or 8 of the 8 signatures predicted the samples to be from high-risk patients), the degree of the survival benefit increased as the GSS decreased (Figure 3D). When

the GSS was greater than 1, patients were predicted to have hazardous effects on survival (Figure 3D), which was in agreement with the results illustrated in Figure 3A-C.

Discussion

In this study, we demonstrated that CSS sets significantly improved the prediction accuracy of prognosis in patients with stage II CRC. The prediction accuracy for low- and high-risk disease significantly outperformed other gene signatures such as Oncotype DX and ColoPrint. In particular, the CSS set-defined high-risk group contained 2 to 3 times more real high-risk samples (55%) than those defined by Oncotype DX (22%)¹⁴ and ColoPrint (22%-26%).^{15,16} Moreover, we showed that the CSS set-defined high-risk group gained significant survival benefits from fluorouracil-based adjuvant therapy. Furthermore, the robustness of the CSS sets was validated in 767 patients with

stage II disease from 11 independent cohorts. We further showed the robustness of the CSS set approach by changing different discovery and validation cohorts (see the Supplement).

Thus far, clinically useful cancer biomarkers remain rare because cancer is a complex disease. However, the complexity of cancer can be more clearly understood by analyzing several distinctive and complementary capabilities (cancer hallmarks or traits) that enable tumor growth and metastasis dissemination.³⁴ Therefore, we propose that CSS sets can be used to more thoroughly capture the complex nature of the disease (Figure 1). Indeed, we showed that CSS sets significantly boosted prediction performance.

Fluorouracil is a first-line drug for colon cancer that has been used to treat patients with stage III disease for the last 20 years. It is expected that high-risk patients with stage II disease could gain survival benefit from fluorouracil adjuvant chemotherapy. If a predicted high-risk group of stage II CRC samples contains too many false positives (ie, many nonrecurred samples), it could be very hard to examine the survival benefit of adjuvant chemotherapy in that predicted high-risk group. The prediction accuracies of Oncotype DX¹⁵ and ColoPrint^{16,17} for high-risk groups of stage II CRC are 22% and 22% to 26%, respectively, suggesting that the predicted high-risk groups contain about 80% of nonrecurred samples. Not surprisingly, both Oncotype DX- and ColoPrint-defined high-risk patients with stage II disease do not gain survival benefit from adjuvant chemotherapy. In this study, the CSS set-defined high-risk group of patients with stage II CRC contained 55% recurred disease. In this case, the CSS set-defined high-risk group of stage II CRC did gain significant survival benefit from adjuvant chemotherapy ($P = .004$). The CSS sets could be used to identify a subset of patients with stage II CRC for receiving fluorouracil-based adjuvant chemotherapy.

We also demonstrated that adjuvant fluorouracil-based chemotherapy in the CSS set-defined low-risk and intermediate-risk patients with stage II disease are more likely to do harm than good. These results point to a potential resolution of the 20-year-old debate of adjuvant chemotherapy in patients with stage II disease. Further successful validations of these results will lead to consensus recommendations for by various professional clinical organizations.

Conclusions

Generally, 20% of patients with stage II CRC experience recurrence within 5 years.^{2,24} Therefore, ideally, a prospective clinical validation trial of 5000 to 6000 such patients could be conducted. Patients' age, sex, location of the tumors, pT1 to pT4 stages, and other clinical features should be used to characterize the clinical nature of the stage II disease. Tumor purity lower than 75% should be excluded from the trial. Frozen tumors or tumor paraffin blocks could be used to extract RNA samples. Gene expression profiling could be conducted using Affymetrix arrays (or designing an Affymetrix-based customized array for the CSS sets). Each sample's data would be meta-normalized with GEO cohorts GSE37892 + GES33113 and then predicted to be from low-, intermediate-, or high-risk patients using the CSS sets. If a sample is predicted as low-risk, no chemotherapy would be administered to that patient. The samples predicted to be from either intermediate- or high-risk patients would be randomized to compare fluorouracil-based adjuvant treatment with no adjuvant treatment within 5 years of follow-up. Experiences in clinical trials⁴¹ for validating breast cancer gene signatures (ie, MammaPrint⁴² or Oncotype DX) could be used to help the design as well.

ARTICLE INFORMATION

Accepted for Publication: July 20, 2015.

Published Online: October 22, 2015.
doi:10.1001/jamaoncol.2015.3413.

Author Contributions: Drs Gao, Tibiche, and Zou contributed equally to this work and had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: O'Connor-McCourt, Wang.

Acquisition, analysis, or interpretation of data: Gao, Tibiche, Zou, Zaman, Trifiro.

Drafting of the manuscript: Wang.

Critical revision of the manuscript for important intellectual content: Gao, Tibiche, Zou, Zaman, Trifiro, O'Connor-McCourt, Wang.

Statistical analysis: Gao, Tibiche, Zou.

Obtained funding: O'Connor-McCourt, Wang.

Administrative, technical, or material support:

Zaman, O'Connor-McCourt, Wang.

Study supervision: Trifiro, Wang.

Conflict of Interest Disclosures: None reported.

Funding/Support: This work was supported by the National Research Council Canada. Dr Gao is supported by a visiting fellowship from the China Scholarship Council.

Role of the Funder/Sponsor: The National Research Council Canada and the China Scholarship Council had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

REFERENCES

- Black ER, Falzon L, Aronson N. Gene Expression Profiling for Predicting Outcomes in Stage II Colon Cancer. 2012. <http://www.ncbi.nlm.nih.gov/books/NBK115808/>. Accessed September 9, 2015.
- Dotan E, Cohen SJ. Challenges in the management of stage II colon cancer. *Semin Oncol*. 2011;38(4):511-520.
- Kelley RK, Venook AP. Prognostic and predictive markers in stage II colon cancer: is there a role for gene expression profiling? *Clin Colorectal Cancer*. 2011;10(2):73-80.
- Meropol NJ. Ongoing challenge of stage II colon cancer. *J Clin Oncol*. 2011;29(25):3346-3348.
- Vachani C, Giantonio B. Stage II colon cancer: To treat or not to treat? 2013. <http://www.oncolink.org/types/article.cfm?c=124&id=9621>.
- Scottish Intercollegiate Guidelines Network. Diagnosis and management of colorectal cancer

guidelines. 2011. <http://www.sign.ac.uk/pdf/sign126.pdf>. Accessed September 10, 2015.

7. ASCO Institute for Quality. ASCO Recommendations on Adjuvant Chemotherapy for Stage II Colon Cancer. <http://www.instituteforquality.org/asco-recommendations-adjuvant-chemotherapy-stage-ii-colon-cancer>. Accessed September 17, 2015.

8. National Comprehensive Cancer Network. NCCN Guidelines for Treatment of Cancer by Site: Colon/Rectal Cancer. http://www.nccn.org/professionals/physician_gls/f_guidelines.asp#site. Accessed September 17, 2015.

9. Fang SH, Efron JE, Berho ME, Wexner SD. Dilemma of stage II colon cancer and decision making for adjuvant chemotherapy. *J Am Coll Surg*. 2014;219(5):1056-1069.

10. Gray R, Barnwell J, McConkey C, Hills RK, Williams NS, Kerr DJ; Quasar Collaborative Group. Adjuvant chemotherapy versus observation in patients with colorectal cancer: a randomised study. *Lancet*. 2007;370(9604):2020-2029.

11. O'Connor ES, Greenblatt DY, LoConte NK, et al. Adjuvant chemotherapy for stage II colon cancer with poor prognostic features. *J Clin Oncol*. 2011;29(25):3381-3388.

12. Li J, Lenferink AE, Deng Y, et al. Identification of high-quality cancer prognostic markers and metastasis network modules. *Nat Commun*. 2010; 1:34.
13. Park YY, Lee SS, Lim JY, et al. Comparison of prognostic genomic predictors in colorectal cancer. *PLoS One*. 2013;8(4):e60778.
14. Gray RG, Quirke P, Handley K, et al. Validation study of a quantitative multigene reverse transcriptase-polymerase chain reaction assay for assessment of recurrence risk in patients with stage II colon cancer. *J Clin Oncol*. 2011;29(35):4611-4619.
15. Kopetz S, Tabernero J, Rosenberg R, et al. Genomic classifier ColoPrint predicts recurrence in stage II colorectal cancer patients more accurately than clinical factors. *Oncologist*. 2015;20(2):127-133.
16. Salazar R, Roepman P, Capella G, et al. Gene expression signature to improve prognosis prediction of stage II and III colorectal cancer. *J Clin Oncol*. 2011;29(1):17-24.
17. Maak M, Simon I, Nitsche U, et al. Independent validation of a prognostic genomic signature (ColoPrint) for patients with stage II colon cancer. *Ann Surg*. 2013;257(6):1053-1058.
18. Srivastava G, Renfro LA, Behrens RJ, et al. Prospective multicenter study of the impact of oncoPrint DX colon cancer assay results on treatment recommendations in stage II colon cancer patients. *Oncologist*. 2014;19(5):492-497.
19. Venook AP, Niedzwiecki D, Lopatin M, et al. Biologic determinants of tumor recurrence in stage II colon cancer: validation study of the 12-gene recurrence score in cancer and leukemia group B (CALGB) 9581. *J Clin Oncol*. 2013;31(14):1775-1781.
20. Birnbaum DJ, Laibe S, Ferrari A, et al; COL2 Project. Expression profiles in stage II colon cancer according to APC gene status. *Transl Oncol*. 2012; 5(2):72-76.
21. de Sousa E Melo F, Colak S, Buikhuisen J, et al. Methylation of cancer-stem-cell-associated Wnt target genes predicts poor prognosis in colorectal cancer patients. *Cell Stem Cell*. 2011;9(5):476-485.
22. Jorissen RN, Gibbs P, Christie M, et al. Metastasis-associated gene expression changes predict poor outcomes in patients with Dukes stage B and C colorectal cancer. *Clin Cancer Res*. 2009;15 (24):7642-7651.
23. Kikuchi A, Ishikawa T, Mogushi K, et al. Identification of NUCKS1 as a colorectal cancer prognostic marker through integrated expression and copy number analysis. *Int J Cancer*. 2013;132 (10):2295-2302.
24. Marisa L, de Reyniès A, Duval A, et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med*. 2013;10(5): e1001453.
25. Reid JF, Gariboldi M, Sokolova V, et al. Integrative approach for prioritizing cancer genes in sporadic colon cancer. *Genes Chromosomes Cancer*. 2009;48(11):953-962.
26. Sheffer M, Bacolod MD, Zuk O, et al. Association of survival and disease progression with chromosomal instability: a genomic exploration of colorectal cancer. *Proc Natl Acad Sci U S A*. 2009; 106(17):7131-7136.
27. Smith JJ, Deane NG, Wu F, et al. Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. *Gastroenterology*. 2010;138(3):958-968.
28. Staub E, Groene J, Heinze M, et al. An expression module of WIPF1-coexpressed genes identifies patients with favorable prognosis in three tumor types. *J Mol Med (Berl)*. 2009;87(6): 633-644.
29. Sveen A, Agesen TH, Nesbakken A, Rognum TO, Lothe RA, Skotheim RI. Transcriptome instability in colorectal cancer identified by exon microarray analyses: associations with splicing factor expression levels and patient survival. *Genome Med*. 2011;3(5):32.
30. Tsukamoto S, Ishikawa T, Iida S, et al. Clinical significance of osteoprotegerin expression in human colorectal cancer. *Clin Cancer Res*. 2011;17 (8):2444-2450.
31. Chen DT, Hernandez JM, Shibata D, et al. Complementary strand microRNAs mediate acquisition of metastatic potential in colonic adenocarcinoma. *J Gastrointest Surg*. 2012;16(5): 905-912.
32. Cui Q, Ma Y, Jaramillo M, et al. A map of human cancer signaling. *Mol Syst Biol*. 2007;3:152.
33. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-674.
34. Wang E, Zaman N, Mcgee S, Milanese JS, Masoudi-Nejad A, O'Connor-McCourt M. Predictive genomics: a cancer hallmark network framework for predicting tumor clinical phenotypes using genome sequencing data. *Semin Cancer Biol*. 2015; 30:4-12.
35. Zaman N, Li L, Jaramillo ML, et al. Signaling network assessment of mutations and copy number variations predict breast cancer subtype-specific drug targets. *Cell Rep*. 2013;5(1): 216-223.
36. Wang E, Lenferink A, O'Connor-McCourt M. Cancer systems biology: exploring cancer-associated genes on cellular networks. *Cell Mol Life Sci*. 2007;64(14):1752-1762.
37. Wang E, Zou J, Zaman N, Beitel LK, Trifiro M, Paliouras M. Cancer systems biology in the genome sequencing era: part 1, dissecting and modeling of tumor clones and their networks. *Semin Cancer Biol*. 2013;23(4):279-285.
38. Wang E, Zou J, Zaman N, Beitel LK, Trifiro M, Paliouras M. Cancer systems biology in the genome sequencing era: part 2, evolutionary dynamics of tumor clonal networks and drug resistance. *Semin Cancer Biol*. 2013;23(4):286-292.
39. Snaebjornsson P, Coupe VM, Jonasson L, Meijer GA, van Grieken NC, Jonasson JG. pT4 stage II and III colon cancers carry the worst prognosis in a nationwide survival analysis: Shepherd's local peritoneal involvement revisited. *Int J Cancer*. 2014; 135(2):467-478.
40. Swamy R. Histopathological reporting of pT4 tumour stage in colorectal carcinomas: dotting the 'i's and crossing the 't's. *J Clin Pathol*. 2010;63(2): 110-115.
41. Cardoso F, Piccart-Gebhart M, Van't Veer L, Rutgers E; TRANSBIG Consortium. The MINDACT trial: the first prospective clinical validation of a genomic tool. *Mol Oncol*. 2007;1(3):246-251.
42. Agendia. MammaPrint 70-gene Breast Cancer Recurrence Assay. <http://www.agendia.com/healthcare-professionals/breast-cancer/mammaprint/>. Accessed September 17, 2015.