The following protocol information is provided solely to describe how the authors conducted the research underlying the published report associated with the following article:

Association between results of a gene expression signature assay and recurrence-free interval in patients with stage II colon cancer on Alliance (CALGB) 9581

DOI: 10.1200/JCO.2015.65.4699

Protocol Update #09
10/15/2014

ALLIANCE FOR CLINICAL TRIALS IN ONCOLOGY

PROTOCOL UPDATE TO CALGB 150705

CORRELATIVE SCIENCE STUDIES IN COLON CANCER
A Companion Study to CALGB 9581 and 89803

[X]	**Update:**	[ ]	**Status Change:**
[ ]	Eligibility changes	[ ]	Activation
[ ]	Therapy / Dose Modifications / Study Calendar changes	[ ]	Closure
[ ]	Informed Consent changes	[ ]	Suspension / temporary closure
[ ]	Scientific / Statistical Considerations changes	[ ]	Reactivation
[ ]	Data Submission / Forms changes		
[X]	Editorial / Administrative changes		
[ ]	Other :		

No IRB approval of this amendment is required. Please follow your local IRB guidelines

UPDATES:

- In keeping with new CTEP PIO requirements, the name of the lead group, "Alliance for Clinical Trials in Oncology," has been added to the title page of the protocol.
- In the bottom left hand corner of the front cover page, the following text has also been added: "Participating Organizations: ALLIANCE / Alliance for Clinical Trials in Oncology."

A replacement protocol document has been issued

ATTACH TO THE FRONT OF EVERY COPY OF THIS PROTOCOL

Activation Date: 07/15/07
Includes Update #09

ALLIANCE FOR CLINICAL TRIALS IN ONCOLOGY

CALGB 150705

CORRELATIVE SCIENCE STUDIES IN COLON CANCER

A Companion Study to CALGB 9581 and 89803

GI Correlative Science Vice Chair
Monica Bertagnolli, MD
Brigham And Women's Hospital
Tel: 617-732-8910
Fax: 617-582-6177
mbertagnolli@partners.org

GI Pathology Cadre Leader
Wendy L. Frankel, MD
The Ohio State University Medical Center
Tel: 814-293-8496
Fax: 614-293-2779
wendy.frankel@osumc.edu

GI Committee Chair
Alan Venook, MD
Univ. of California, San Francisco
Tel: 415-353-2745
Fax: 415-353-9959
venook@cc.ucsf.edu

Faculty Statisticians

Donna Niedzwicki, PhD
Tel: 919-681-5030
Fax: 919-681-8028
niedz001@mc.duke.edu

Kouros Owzar, PhD
Tel: 919-681-1829
Fax: 919-681-8028
kouros.owzar@duke.edu

Staff Statistician
Xing (Cynthia) Xe, MS
Tel: 919-681-2044
Fax: 919-681-8028
xing.ye@duke.edu

Data Coordinator
Kathe Douglas
Tel: 919-668-9368
Fax: 919-668-9348
Kathe.Douglas@duke.edu

Protocol Coordinator
Shivani Shah
Tel: 773-834-4518
Fax: 312-345-0117
sshah11@uchicago.edu

Participating Organizations
ALLIANCE / Alliance for Clinical Trials in Oncology

CO-CHAIRS

SECTION B1
A PROSPECTIVE STUDY OF DIET, LIFESTYLE FACTORS, AND MEDICATIONS AND INTERACTIONS WITH MOLECULAR MARKERS AMONG PATIENTS WITH STAGE III COLON CANCER

Charles Fuchs, MD, MPH – 617-632-5840
cfuchs@partners.org

SECTION B2
GENOMIC PREDICTORS OF SURVIVAL AND RESPONSE TO THERAPY IN STAGE III COLON CANCER PATIENTS RECEIVING ADJUVANT THERAPY FOLLOWING RESECTION

Frederic Waldman, MD, PhD – 416-476-3821
waldman@cc.ucsf.e*du*

SECTION B3
VALIDATION OF A PROGNOSTIC TEST FOR STAGE II COLORECTAL CANCER:
A RETROSPECTIVE STUDY OF TISSUES FROM CALGB PROTOCOL 9581

Wendy Frankel, MD,– 614-293-7625
Wendy.Frankel@osumc.edu

SECTION B4
A LABORATORY STUDY TO CLINICALLY VALIDATE THE RELATIONSHIP BETWEEN GENOMIC TUMOR EXPRESSION PROFILES AND THE LIKELIHOOD OF RECURRENCE IN PATIENTS WITH RESECTED STAGE II COLON CANCER

Robert Warren, MD – 415-353-9291
warrenr@surgery.ucsf.edu

SECTION B5
EPIGENETIC PREDICTORS OF SURVIVAL AND RESPONSE TO THERAPY IN STAGE III COLON CANCER PATIENTS RECEIVING ADJUVANT THERAPY FOLLOWING RESECTION THIS STUDY USES SAMPLES FROM CALGB 89803.

William Grady, MD – 206-667-1107
wgrady@fhcrc.org

SECTION B6
NEW PROGNOSTIC BIOMARKERS IN COLON CANCER: A STUDY USING TISSUE FROM CALGB 89803.

Shuji Ogino, MD, PhD – 617-632-3978
sogino@partners.org

CALGB Central Office
230 West Monroe Street, Suite 2050
Chicago, IL 60606-4703
Phone: 773-702-9171
Fax: 312-345-0117
www.calgb.org

CALGB Statistical Center
Hock Plaza
2424 Erwin Road, Suite 802
Durham, NC 27705
Tel: 919-668-9348
Fax: 919-668-9350
Biostatistics Fax: 919-681-8028

CALGB Pathology Coordinating Office
The Ohio State University
Innovation Centre
2001 Polaris Parkway
Columbus, OH 43240
Phone: 614-293-7073
Fax: 614-293-7967
path.calgb@osumc.edu

CORRELATIVE SCIENCE STUDIES IN COLON CANCER
A Companion Study to CALGB 9581 and 89803

SCHEMA

ELIGIBILITY:

- Patient must have been registered to CALGB 9581 or 89803.
- A sufficient number of tissue samples must be available at the CALGB Pathology Coordinating Office.

METHODS:

- **Treatment studies CALGB 9581 and 89803**: Blocks and/or slides have been submitted from cooperative group institutions participating in the treatment studies to the respective cooperative group repository.
- Samples collected from patients enrolled on CALGB 9581 and 89803 should now be stored at the CALGB Pathology Coordinating Office.
- **Tissue samples or isolated DNA** will then be sent to individual investigators as follows:

		From Treatment Study:
Section B1 (K-ras, Cox-2)	• 1 x 15 µm slide for DNA and • 3 x 5 µm slides for IHC forwarded to Shuji Ogino, M.D.	89803
Section B2 (MSI)	• 1 x 100 ng aliquot of isolated tumor DNA forwarded to Frederic Waldman, M.D., Ph.D	89803
Section B3	• 1 x 4 µm slide for H&E review • 2 x 10 µm slides for array analysis forwarded to Almac Diagnositics	9581
Section B4	• 3 x 10 µm slides for PCR analysis forwarded to Genomic Health	9581
Section B5	• TMA slides for IHC and • 1 x 100 ng aliquot of isolated tumor DNA forwarded to William Grady, M.D.	89803
Section B6	• 1 x 5 µm slide for H&E review 2 x 5 µm slides for IHC • 1 x 15 µm slide for sequencing forwarded to Shuji Ogino, M.D., Ph.D.	89803

- **Investigators will perform analyses** per the procedures described in Section B and send coded data to the CALGB Statistical Center.

REGISTRATION PROCEDURES: There will be no additional registration from individual institutions for this study.

INSTITUTIONAL REVIEW BOARD: Review and approval at the institution where the laboratory work will be performed is required.

Table of Contents

SECTION A INTRODUCTION/BACKGROUND, ELIGIBILITY, IDENTIFICATION OF REQUIRED SAMPLES

1.0 INTRODUCTION AND BACKGROUND

1.1 Study Description

This protocol is divided into sections A and B.

Section A of this protocol describes the background and overall goals of the correlative science studies in colon cancer to be conducted using the samples previously collected on CALGB Studies 9581 and 89803. In addition, the samples available for use are identified and the procedures for their dispersement are described. Section B is divided into subsections, which describe the laboratory studies. These subsections may be amended as study objectives are met or as new technologies are developed.

This protocol describes laboratory studies performed on tissue samples from colon cancer patients. These samples are provided by investigators at cooperative group institutions who intended to treat their eligible patients on either of the following CALGB trials: CALGB 9581 or 89803. The institutional investigators have sent these samples either directly to the their cooperative group's pathology coordinating office or to the CALGB Pathology Coordinating Office for subsequent distribution.

1.2 Background

In 2006, an estimated 156,000 new cases of colorectal cancer are expected in the United States (1). Surgery is the primary modality of management for these tumors and a resection with 'curative intent' is possible in ~85% of patients. Despite this, more than 60,000 individuals are expected to die from the disease, usually from distant metastases. For patients at high risk of disease recurrence, adjuvant therapies significantly improve survival (2,3). Ideally, only patients who have occult metastases should receive adjuvant therapy because the drugs are toxic, expensive, and only moderately effective against advanced disease.

Pathologic stage criteria do not precisely define which patients are at high risk of disease recurrence following treatment of the primary tumor: 70-80% of stage II and 60-65% of stage III patients receiving current adjuvant chemotherapy will not develop metastases and will be cured with surgery alone. It is important to more precisely identify those patients at greatest risk for recurrence, both to identify those most likely to benefit from adjuvant therapy, and to select those for which adjuvant therapy is likely to be ineffective. In addition to allowing better disease management, an accurate prediction of disease outcome facilitates development of new therapies for patients who fail to respond to current management.

The current criteria used to evaluate prognosis are stage-related and histologic parameters. Stage-related variables measure the extent of disease (depth of invasion of bowel wall, presence or absence of metastasis in regional nodes or distant sites, number of involved nodes). Histologic variables assess the state of differentiation of a neoplasm (grade), and pre-operative serum carcinoembryonic antigen (CEA) level (4). Most treatment decisions are made using the TNM staging system (4), which assesses depth of invasion of the bowel wall and tumor status of regional lymph nodes. Unfortunately, these staging systems create categories too large to be useful prognosticators for individual patients and can be applied only after pathologic examination of the resected primary tumor and lymph nodes. This has led various investigators to examine additional genotypic or phenotypic characteristics in colon carcinoma that may be of prognostic value.

In two large adjuvant therapy trials, CALGB 9581 and 89803, analyses are currently underway to determine whether the expression of markers in primary tumors are associated with relapse-

free survival and overall survival. CALGB 9581 randomized a total of 1738 patients to receive either placebo, or monoclonal antibody 17-1A for adjuvant therapy following resection of a stage II colon cancer. CALGB 89803 studied the effect of 5-FU + leucovorin, with or without irinotecan, in adjuvant treatment of 1260 patients following resection of stage III colon cancer. To provide a thorough examination of colorectal tumors, a "panel" approach was developed to study specific tumor cell characteristics such as cell cycle regulators (e.g., p53, p21, p27), angiogenesis (e.g.,VEGF, MVD), chemotherapy response (e.g., thymidylate synthase), DNA replication integrity (e.g., microsatellite instability), and "other" potential markers (e.g., 18qLOH). Each of these markers has demonstrated promising preliminary data in single institution settings.

The work conducted under CALGB Protocols 9581 and 89803 resulted in creation of a tissue bank containing tumor and normal tissue samples from a large number of patients with stage II and III colon cancer, uniformly treated under adjuvant therapy protocols. This tissue bank can be used to identify or validate additional markers of treatment outcome or prognosis for this common disease. In addition, additional work conducted with this resource allows a comparison of the relative contributions of each marker and marker category to outcome assessment. The purpose of this protocol, CALGB 150705, is to enable additional correlative science studies using the tissue resources collected under CALGB 9581 and 89803.

2.0 ELIGIBILITY CRITERIA

2.1 Registration to CALGB 9581 or 89803.

2.2 Samples present within the CALGB PCO or at the Institutions providing treatment that are sufficient to meet study aims.

2.3 Institutional Review Board (IRB) review and approval at the institution where the laboratory work will be performed is required.

2.4 Informed Consent: CALGB does not require that a separate consent form be signed for this study.

2.4.1 The subject population to be studied in this protocol includes patients selected from either of the following CALGB treatment protocols: CALGB 9581 or 89803. All such patients have signed (or will sign) a written informed consent document meeting all federal, state, and institutional guidelines as part of entry into those trials.

2.4.2 All samples to be studied are obtained and stored as part of the patient's respective treatment trial. The material and data obtained from the patient's protocol record will be used to obtain appropriate clinical information. In no instance will the patient be contacted directly.

2.4.3 There should be no physical, psychological, social, or legal risks associated with this study. No invasive procedures are recommended or requested.

2.4.4 All appropriate and necessary procedures will be utilized to maintain confidentiality. All patients who have had samples submitted for analysis will have their CALGB study number used to identify specimens.

2.4.5 This study does not require direct patient contact and no specific risk or benefits to individuals involved in the trial are anticipated. It is likely, however, that the information gained will substantially help similar patients in the future.

3.0 IDENTIFICATION OF REQUIRED SAMPLES

3.1 Samples prospectively collected from CALGB 9581 and 89803

Tissue samples previously collected from CALGB 9581 or 89803 were sent directly to the appropriate cooperative group repositories or the CALGB PCO. These samples should now be stored at the CALGB Pathology Coordinating Office. Appropriate samples are to be identified and the patient CALGB ID number recorded and registered to the CALGB database.

Samples used for the studies in this protocol will be dispersed by the CALGB PCO. All specimens collected under CALGB 9581 or 89803 not already present in the CALGB PCO, will be transferred to this location before distribution to research laboratories.

3.2 Data Collection/Management Procedures

Clinical data for patients treated under CALGB 9581 and 89803, including patient characteristics, pretreatment laboratory data, and treatment are available in the CALGB Statistical Center. The data obtained laboratory analyses performed under this protocol will be sent to CALGB Statistical Center and entered into the database on a quarterly basis.

4.0 REFERENCES

1. Howe HL, Wu X, Ries LAG, Cokkinides V, et al. Annual report to the natin on the status of cancer, 1975-2003, featuring cancer among US Hispanic/latino populations. Cancer 2006; 107(8):1711-1742.
2. Benson AB 3rd. New approaches to the adjuvant therapy of colon cancer. Oncologist 2006; 11(9):973-980.
3. Chau I, Cunningham D. Adjuvant therapy in colon cancer – what, when and how?. Ann Oncol 2006; 17(9):1347-1359.
4. Kehoe J, Khatri VP. Staging and prognosis of colon cancer. Surg Oncol Clin N Am 2006; 15(1):129-146.

SECTION B3: VALIDATION OF A PROGNOSTIC TEST FOR STAGE II COLORECTAL CANCER:

A RETROSPECTIVE STUDY OF TISSUES FROM CALGB PROTOCOL 9581

This study uses samples from CALGB 9581.

Co-Investigators: Wendy Frankel, MD
Wendy.Frankel@osumc.edu
Naimish Pandya, MD
npandya@umm.edu

1.0 INTRODUCTION/BACKGROUND

The worldwide incidence of stage II colorectal cancer (AJCC/UICC TNM staging: T3N0M0 and T4N0M0) is approximately 300,000 cases per year [1, 2], of which, despite surgical resection of all detectable tumor, 20-25% of patients will develop recurrent disease/relapse post surgery without additional treatment [3]. These patients are likely to benefit from adjuvant therapies. However, if all stage II patients received adjuvant chemotherapy, 75-80% would be treated without benefit. Unfortunately, present tumor staging methods do not allow us to accurately identify which 20-25% of stage II colorectal cancer patients will suffer relapse.

Tumor stage is determined by histopathologic examination of the tissues removed at surgery, in addition to radiographic assessment of distant organs such as the liver and lungs. However, a number of different approaches to improved colorectal cancer staging are currently under investigation, many involving molecular characterization of the primary tumor. Unfortunately, however, for patients with stage II colorectal cancer no single clinical trial or study to date has conclusively demonstrated a role for molecular markers in the assignment of relapse risk and hence likely benefit from adjuvant therapy.

There is therefore a clear clinical need to identify robust prognostic molecular markers in order to improve the clinical management of the stage II colorectal cancer patient population, by identification of the 20-25% at high risk of relapse that may benefit from additional adjuvant therapy.

Almac Diagnostics has developed a gene expression signature capable of stratifying stage II colorectal patients into patients at high risk and patients at low risk of relapse. In order to gain approval for the test, Almac Diagnostics must perform a clinical validation of the prognostic gene signature. The CALGB 9581 study contains a large enough cohort of stage II colon cancer patients for such a validation study.

1.1 Development of Assay

The Almac Diagnostics gene expression signature was developed in two stages. The first involved creation of a Colorectal Cancer Disease Specific Array (CCDSA) that included genes differentially expressed in human colorectal cancers. Once this was accomplished, then a set of human stage II colorectal cancers for which clinical outcome data were known were used to refine the CCDSA into a signature specific for "high risk", i.e. recurrent disease. This assay is known as the "prognostic signature". The methods to develop the prognostic signature are described below.

The CCDSA content was generated by a combination of high throughput in-house sequencing, public database mining and experimental investigation. It was optimized to enable profiling of RNA extracted from fresh and paraffin-embedded tissue samples.

The CCDSA contains 61,528 probesets covering 52,306 transcripts expressed in colorectal tissues, of which:

- 21,968 transcripts (42%) are present in the human RefSeq database
- 26,676 transcripts (51%) are not present in the human RefSeq database
- 7% of the content represents expressed anti-sense transcripts to annotated genes
- 20,000 transcripts are not detected by current generic microarray platforms

Sequence content analysis confirms that the CCDSA currently offers the most comprehensive platform available for the study of colorectal cancer.

Reproducibility and stability of the Colorectal Cancer DSA (CCDSA) were assessed by profiling a series of technical replicates on both the CCDSA and the Affymetrix Plus2 Human Genome Array. The experimental design of the technical assessment study can be seen in Table 9. All arrays were normalized using DChip using Invariant Set Normalization. The reproducibility and stability of the data were assessed by analyzing the Spearman Rank Correlation between the samples in each wing of the experiment. Initially a "whole array" correlation was performed (assessing the consistency of data across the array including background noise). The second measure looked at the correlation of the expression data for those called "present" by the Affymetrix MAS 5.0 algorithm (detection p-value <0.05) and expressed above the background level.

Table 9. Experimental design of the technical assessment study

<table>
<tr><th>RNA Later</th><th>FFPE</th><th></th><td rowspan="5">RNA was extracted from a "quad set" of matched tissues from a single stage II CRC patient, consisting of:

-RNA-later preserved Normal tissue
-RNA-later preserved Tumor tissue
-FFPE Normal tissue
-FFPE Tumor tissue</td></tr>
<tr><td>Tumor N = 5</td><td>Tumor N = 5</td><td rowspan="2">Almac Diagnostics CCDSA</td></tr>
<tr><td>Normal N = 5</td><td>Normal N = 5</td></tr>
<tr><td>Tumor N = 5</td><td>Tumor N = 5</td><td rowspan="2">Affymetrix HG-U133 Plus 2</td></tr>
<tr><td>Normal N = 5</td><td>Normal N = 5</td></tr>
</table>

The transcripts detected by profiling the matched RNA-later/FFPE tissue samples on the CCDSA were compared. Present calls were calculated based on the Affymetrix 5.0 "Present Call" algorithm and an expression level above background. The results of this analysis are shown in Table 10.

Table 10. Retention of data generated from RNA-later and FFPE

	Transcipts Detected in RNA-Later	Transcripts Detected in FFPE	No. of FFPE Detected Transcripts also Detected in RNA-Later	Percent FFPE-Detected Transcripts also Detected in RNA-Later
Almac CCDSA	23,869	18,723	16,415	88%

As a comparison between FFPE and RNAlater samples Table 10 shows that 88% of transcripts detected in FFPE were detected in RNAlater. This high level of retention effectively creates the conditions for good 'correlation' between FFPE and RNAlater. Table 10 also shows that

23,869 transcripts were detected in RNAlater while 16,415 transcripts were detected in FFPE and RNA later. This means that 70% of transcripts detected in RNAlater were also detected in FFPE.

1.2 Generation of the Prognostic Signature using Stage II Colorectal FFPE Tissue

After receiving IRB or Ethics Committee approval, tumor tissue specimens were collected from 359 patients with stage II colorectal at US and European centers between 1990 and 2007. The patients had not received pre-operative or post-operative radiotherapy or chemotherapy. The histopathology of each specimen was reviewed to confirm diagnosis of stage II and tumor involvement.

Of the 359 patient specimens, 227 (63%) were classified as low risk of relapse (good prognosis) on the basis of five year disease-free survival post-resection of the primary tumor. 132 (37%) of the specimens were classified as high risk of relapse (poor prognosis) on the basis of relapse (cancer recurrence) within five years of surgical resection of the primary tumor. In this study there was a deliberate enrichment of high risk patients relative to the general stage II population in which 20-25% of patients recur.

Total RNA was extracted from a single 10 micron FFPE section (minimum 50% tumor cell content) cut from each of the 359 tissue specimens, using an in-house protocol and commercially available RNA extraction kit. The 359 FFPE patient samples were randomly separated into a training set and an independent test set as follows:

Training set (60%): n=215 (142 good and 73 poor prognosis cases)

Test set (40%): n=144 (85 good and 59 poor prognosis cases)

Biotinylated aRNA targets were generated using GeneChip® Two Cycle Target Labeling and Control Reagents (Affymetrix) and hybridized to Colorectal Cancer Disease Specific Arrays (CCDSAs), which were then scanned using the Affymetrix GeneChip® 7G Scanner.

Array signal processing was performed as follows: array background signal correction was performed with the Robust Multi-array Average (RMA) procedure applied in a reference context (RefRMA). Array signal processing was performed on the training dataset within each level of cross-validation with model parameters (reference quantile and probe effects for summarization) saved for application to subsequent test dataset arrays. Gene selection was performed by iterative feature elimination within cross-validation. Classification models were developed on the training set portions by partial least squares discriminant analysis (PLS-DA). The optimal number of prognostic transcripts, feature selection and estimates for the final performance of the final signature, was determined using ten repeats of 5-fold cross-validation.

1.3 Preliminary Results

Based upon the Affymetrix GeneChip technology, the CCDSA provides multiple independent measurements for each transcript ensuring high quality and reproducibility of the gene expression data. Table 11 contains correlation coefficients and coefficients of variance for technical replicates run on the Affymetrix Plus 2.0 and CCDSA arrays using FFPE and RNAlater samples. Effectively, the tables show that data generated on the CCDSA are highly reliable and reproducible. Values are virtually identical between the Affymetrix Plus 2.0 and the CCDSA; therefore the data produced by the DSA are equivalent to that of the Affymetrix Plus 2.0 in terms of reliability and reproducibility. RNAlater and FFPE technical replicates demonstrate 99.6% and 99.2% whole array correlation respectively on the Almac CCDSA array.

Table 11. Assessment of the technical reproducibility of the Almac CCDSA

	Tissue Storage Method	Whole Array Correlation (%)	"Exposed Genes" Correlation (%)	Co-efficient of Variation
Almac CCDSA	RNA Later	99.6	99.6	5.81
	FFPE	99.2	99.0	10.43
Affymetrix HG-U133 Plus 2	RNA Later	99.7	99.6	10.72
	FFPE	99.3	99.2	11.36

A 634 probeset prognostic signature was derived from the gene expression data of the 215 training set samples. Good separation between prognostic groups and a high concordance between training and test set data are indicated by the display of the first two latent variables on PLS-DA (data not shown).

175 of the 634 probesets (27.6%) selected were found to be unique to the CCDSA. In addition, 151 of the 634 probesets (23.8%) were expressed in their reverse orientation as endogenous anti-sense transcripts. On the independent test set of 144 patients, the 634 probesets prognostic signature displayed an overall classification accuracy of 62.8%, with associated sensitivity and specificity of 52.7% and 70.3%, respectively. Based on these estimates of sensitivity and specificity, positive (PPV) and negative (NPV) predictive values were estimated for a general stage II patient population assuming an underlying prevalence of 25% poor prognosis patients and 75% good prognosis patients. The estimated PPV and NPV are 37.2% and 81.7%, respectively.

Cox proportional hazards regression analysis of the good and poor prognosis subgroups identified by the prognostic signature in the independent test set demonstrated a hazard ratio of 2.526 (p=0.0030), indicating a higher risk of and significantly shorter time to cancer recurrence among patients predicted as having a poor prognosis, compared to those predicted as having a good prognosis.

A receiver operating characteristic (ROC) curve was generated on the training set, with an area under the curve (AUC) performance characteristic of 0.68, indicating that the prognostic signature had a high predictive power for stage II colorectal cancer patient outcome.

The performance of the 634 probeset prognostic signature was compared against current standard pathological parameters used in evaluating prognosis in stage II colorectal cancer patients, including: patient age, tumor stage, tumor grade and the number of regional lymph nodes retrieved and assessed. Multivariate Cox proportional hazards regression analysis revealed the 634 probeset prognostic signature (Hazard ratio, HR: 2.577, p=0.007). Patient Age (HR: 1.045, p=0.0042) and Tumor location (proximal vs distal) (HR: 2.154, p=0.0081) as the only statistically significant independent prognostic factors (Table 12).

Table 12. Comparison of CCDSA to standard pathological parameters (based on Goldberg et al., 2004)

	Univariate			Multivariate		
	HR	CI	p	HR	CI	p
Tumor Stage (T4 vs T3)	1.230	(0.667-2.269)	0.5067	1.614	(0.839-3.101)	0.1512
Patient Age	1.039	(1.010-1.069)	0.0086	1.045	(1.014-1.077)	0.0042
Tumor Grade (I,II,III)	0.875	(0.506-1.511)	0.6312	0.662	(0.354-1.239)	0.1968
Tumor Location (Proximal vs Distal)	1.766	(1.075-2.901)	0.0248	2.145	(1.220-3.774)	0.0081
Gender	1.165	(0.713-1.901)	0.5426	0.961	(0.546-1.693)	0.8917
Mucinous subtype	0.825	(0.418-1.627)	0.5787	0.927	(0.456-1.883)	0.8336
No. of Nodes Retrieved	1.007	(0.983-1.032)	0.5678	1.014	(0.989-1.040)	0.2796
Prognostic Signature	2.526	(1.536-4.154)	0.0003	2.577	(1.489-4.459)	0.0007

The fact that the prognostic signature HR of 2.577 obtained on multivariate analysis is close to what was observed on univariate analysis (HR: 2.526) implies that the prognostic value of the gene signature is independent of all other clinico-pathologic variables analyzed.

The Almac Diagnostics stage II recurrence classifier consists of 634 probesets The 634 transcripts were analyzed for links with known biological processes through the use of Gene Ontology software. (Table 13).

Table 13. Classification of transcripts with biological process

Ingenuity Canonical Pathways	p-value	Molecules
IGF-1 Signalling	0.0011	PTK2, JUN, CTGF, IGF1, FOXO3, PRKAR2A, RASA1, IGFBP2
PXR/RXR Activation	0.0055	PPARA, SCD, FOXO3, PRKAR2A, INSR, CYP3A5
TGF-β Signalling	0.0081	JUN, BMP2, SMURF2, VDR, SERPINE1, SMURF1
Estrogen Receptor Signalling	0.0089	PRKDC, POLR2A, MED1, TAF15, NCOR1, GTF2F1, MED12, SMARCA4
DNA Methylation and Transcriptional Repression Signalling	0.0100	CHD3, MTA2, RBBP4
Aryl Hydrocarbon Receptor Signalling	0.0120	CTSD, JUN, NFIX, CCND2, MED1,GSTO2 , CYP1B1, SMARCA4
Glucocorticoid Receptor Signalling	0.0138	ICAM1, POLR2A, JUN, HMGB1 (includes EG: 3146), MED1, TAF15, NCOR1, NR3C2, GTF2F1, STAT1, SERPINE1, SMARCA4
HMGB1 Signalling	0.0209	ICAM1, JUN, RHOQ, HMGB1 (includes EG:3146),RHOJ,SERPINE1
Molecular Mechanisms of Cancer	0.0219	PRKDC, BMP2, PRKAR2A, RHOJ, HIF1A, RALBP1, PTK2, JUN, CCND2, RHOQ, RABIF, FANCD2, ARHGEF2, RASA1
Chemokine Signalling	0.0224	PTK2, CALM3, JUN, CAMK1D, LIMK2
Purine Metabolism	0.0257	PKM2, TJP2, HSPD1, RALBP1, DLG3, SMARCA4, INO80, POLR2A, PICK1, POLR3A, BAT1, ATP5G1, NT5E
Sonic Hedgehog Signalling	0.0275	HKR1, ARRB2, PRKAR2A
Semaphorin Signalling in Neurons	0.0302	PTK2, RHOQ, RHOJ, LIMK2
RhoA Signalling	0.0355	PTK2, IGF1, ARHGAP4, ARHGAP12, LIMK2, ARHGAP8

2.0 OBJECTIVE

The objective of this study is to assess the ability of a prognostic gene expression signature to stratify stage II colorectal cancer patients into those who will experience relapse within five years post-surgery (high risk) and those who will experience five-year disease free survival (low risk), without additional treatment.

3.0 METHODS

CALGB statisticians will identify eligible patients treated on C9581 with available samples based on the inventory provided by the CALGB PCO. CALGB statisticians will provide a list of samples that will be uniquely coded by the CALGB PCO and sent to Almac Diagnostics in a blinded fashion for gene expression profiling.

3.1 Sample requirements

One 4 µm and two 10 µm sections of tumor tissue will be required for this study. The 4 µm section will be used for histopathological assessment of tumor content by an independent pathologist. The adjacent 10 µM samples will be used for RNA extraction. The two 10µm samples will be placed within microcentrifuge tubes (1 section per tube) and the 4 µm sample will be placed on a slide and stained with haematoxylin & eosin (H&E) and transported to Almac Diagnostics on the day following sectioning. Samples should be received by Almac Diagnostics within 72 hours of sectioning.

3.2 Gene Expression Profiling

Each FFPE tissue section will be processed using Almac Diagnostics' quality controlled standard operating procedures to determine the stage II Colon Cancer Prognostic Signature.

3.3 Bioinformatic Analyses

Each sample in the validation set will be processed on a one-by-one basis and the quality control assessment (QC) will be carried out on each individual sample. Classification will be evaluated only on arrays that pass QC. The gene signature was developed by pre-processing the microarray data with RefRMA. In the predictive setting, the Affymetrix probe-level data in the validation set will be normalized and summarized to the 215 training set samples and the gene expression indices will be computed using the stored parameters derived from the training set. The PLS classifier will then be applied to the a priori chosen 634 most informative genes by forming the signature-defined weighted sum of their gene expression indices and comparing this compound score with a pre-computed threshold. Almac will be completely blinded to the clinical outcome of the individual validation set samples.

4.0 STATISTICAL CONSIDERATIONS (ORIGINAL)

Data from patients with available tissue samples randomized to CALGB 9581 who meet the study eligibility criteria will be used to independently establish the prognostic value of the gene expression signature. The primary endpoint will be recurrence-free interval (RFI) measured from the date of surgical resection until documented distant recurrence of primary disease or death due to primary disease. Patients with documented new primaries, who died without recurrence or disease related death, or were lost to follow-up are censored at the time of death or last follow-up time. A case-cohort design will be implemented [4].

CALGB 9581 was a randomized trial of MoAb 17-1A (edrecolomab) versus observation in patients with stage II colon cancer. [6]

One thousand seven hundred and thirty-eight patients (n=1,738) were randomized on this trial between May 1998 and May 2002, 865 on the treatment arm and 873 on the observation arm. Primary tissue samples were collected on a subset of patients.

The primary clinical objective of CALGB 9581 was to determine whether adjuvant treatment with MoAb 17-1A improved overall (OS) in patients who have had resection of a stage II (pT3N0 or pT4aN0) colon cancer. OS was measured from study entry until death from any cause. Disease-free survival (DFS) measured from trial entry until documented recurrence of primary disease or death from any cause was studied as a secondary endpoint.

Patients were randomized to either treatment with MoAb 17-1A or observation only with stratification according to degree of differentiation, vascular or lymphatic invasion, and preoperative serum CEA. Therapy consisted of an initial 2-hour infusion of 500 mg of MoAb 17-1A (cycle 1) followed by a lower dose (100 mg) every 28 days for four doses.

Study results were released by the CALGB Data and Safety Monitoring Board after futility boundaries for OS and DFS were met at the fifth planned interim analysis in July 2003. As of December 2009, there were 29.4% patients with a maximum Grade 3 toxicity, 5.8% patients with maximum Grade 4 toxicity, and no deaths attributed to treatment. Median follow-up among surviving patients was 7.9 years on both MoAB17-1A and Observation. Two hundred seventy (270) failures (recurrence of primary disease or death due to primary disease) and 150 disease-related deaths have been observed (n=1738 patients). These comprised 141 failures and 75 disease-related deaths on the MoAB 17-1A arm and 129 failures and 75 disease-related deaths on the observation arm. Median OS and DFS have not been reached. There were no significant differences between treatment arms (OS p=0.71; DFS p=0.64). It was concluded that MoAb 17-1A does not appear to prolong OS or DFS in patients with Stage II colon cancer following resection [5, 6].

Inclusion Criteria

Criteria for inclusion in the current study sample are as follows.

Histopathological Characteristics

- Histopathological proof of adenocarcinoma of the colon only.
- No gross or microscopic evidence of residual disease; all margins, proximal, distal, and radial, must be negative.
- Primary diagnosis of stage II colon cancer according to the AJCC/UICC T and M classification (T3M0 or T4M0 only) and patients with no positive nodes. Patients are not excluded based solely on M stage being unknown.

Therapy

- Surgery of the colon with curative intent only.
- No pre-operative therapy of any kind administered within 1 year of surgery.
- No post-operative therapy administered, except protocol-related edrecolomab in the case of CALGB 9581 samples. Patients in the CALGB 9581 trial who received non-protocol-related therapy (chemotherapy or radiotherapy) after diagnosis of cancer recurrence are eligible for the study.

Each of the inclusion criteria must be met in order for the patient and tumor sample to be considered eligible for the Almac validation study.

Exclusion Criteria

If any of the following exclusion criteria apply, the patient and tumor sample will be considered ineligible for the study.

Histopathological Characteristics

- Non-adenocarcinoma tumor types of the colon.
- Primary tumors other than stage II colon cancers, according to the AJCC/ UICC T and M classification. This means that T1 and T2; M1 and M2 tumor classifications and patients with positive nodes are excluded, indicating nodal involvement and evidence of distant metastases.

Therapy

- Patients who received pre-operative therapy of any kind within 1 year of surgery.
- Patients who did not recur and received non-protocol-related therapy are excluded. Note: patients who received non-protocol therapy at the time of cancer recurrence are eligible.

Sample Size and Power Estimation

One thousand four hundred and fifty-four patients (n=1,454) of the 1,738 patients enrolled on CALGB 9581 meet the study entry criteria. Of these, 902 patients have available tumor samples (Figure 1).

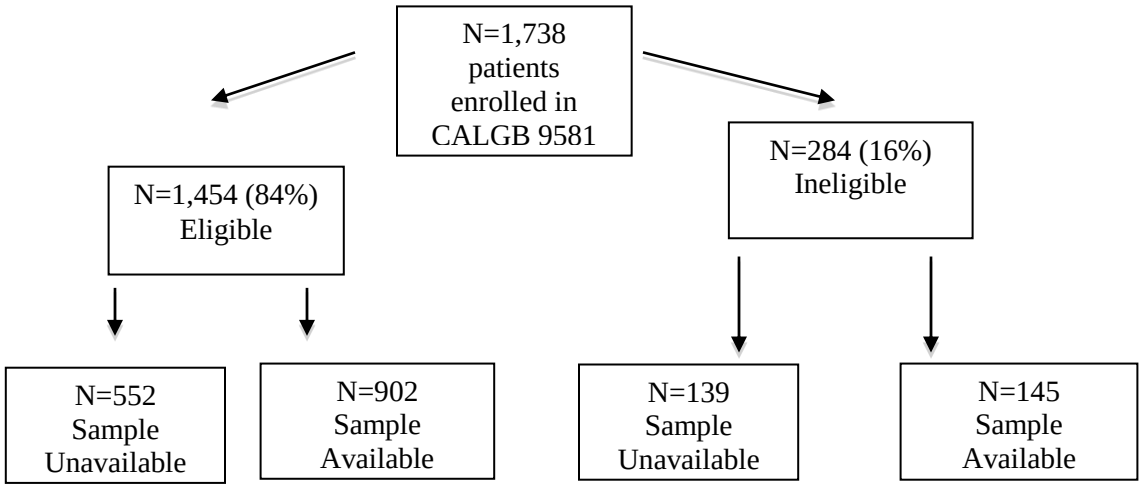

Figure 1. Eligible patients and available sample numbers from CALGB 9581.

For the 902 CALGB 9581 eligible patients with available samples, the minimum follow up time is 2.9 months with a maximum of 12.1 years. The incidence rate is 0.020 recurrences/per person-year. There are 124 (13.8%) events and 778 (86.2%) censored observations. The reverse Kaplan-Meier median follow-up was 8.1 years (95% CI, 7.9 – 8.2).

The impact of selection bias was assessed by comparing the patient characteristics of the 902 CALGB 9581 patients with samples available and the 552 with no available sample. No significant differences were found (data not shown).

A case-cohort design will be used to compare recurrence-free interval (RFI) between high risk and low risk patients as determined by the gene expression signature. RFI is measured from study entry until recurrence of primary cancer or death due to primary cancer. Based on the application of laboratory and microarray quality control procedures in a number of studies, we anticipate a maximum attrition rate of 10%. A sub-cohort of 514 patients will be randomly sampled from the 902 eligible patients with available samples. Random selection will be stratified on treatment (observation; edrecolomab), though survival in the parent study (C9581) did not differ between treatment groups. With 514 of 902 patients sampled in the sub-cohort (a sampling fraction of 0.5698) and the following assumptions: the proportion of samples "marker positive,"p1, is 0.20; the proportion of events in the full cohort, pD, is 0.138; 2-sided α=0.05, a hazard ratio of 1.95 can be detected with approximately 80% power. The sample size of the sub-cohort and associated power were determined based on the methods proposed by Cai and Zeng [7].

Based on the event rate of 0.138 (124/902), we expect 71 events in the sub-cohort and 53 (124-71) events outside the sub-cohort. Of the 53 events outside the sub-cohort we expect that a minimum of 90% (d=48) will be usable. The final size of the analysis data set is, thus, expected to comprise approximately 510 patients, (514+53)*0.9, including 112 events.

Each patient will be classified by the gene expression signature according to a pre-specified threshold as poor (high risk) or good prognosis (low risk). The primary hypothesis that this classification is significantly associated with RFI will be tested using a weighted Cox proportional

hazards model adjusting for standard prognostic factors including age at resection, sex, depth of tumor invasion, tumor grade, tumor location, number of nodes examined, and mismatch repair status (MMR) using a method proposed by Chen and Lo. [8-9] MMR status will be determined based on MLH1 and MSH2 as assessed by immunohistochemistry (i.e., as MMR deficient if either MLH1 or MSH2 is deficient). All prognostic factors were determined at baseline. The detectable hazard ratio under the alternative hypothesis in this study is 1.95, lower than that previously observed.

Univariate Analyses

Univariate analysis of the stage II colon cancer genomic classifier will use the Kaplan Meier product-limit estimator to display the time to event curves for the endpoint. As recommended by the REMARK guidelines, similar analyses will be presented for the other variables being investigated [10].

The estimated effect as a hazard ratio using survival analysis regression will be presented and the likelihood ratio p-value will be reported. The assumptions of the analyses will be evaluated either graphically or with an appropriate statistical test.

Multivariable Analysis

For the analysis of a prognostic marker study, multivariable regression models are the method of choice [11]. Therefore, as stated above, the acceptance of the stage II colon cancer genomic classifier will be determined by multivariable analysis. The multivariable analysis is the primary analysis of the validation study. The covariates are repeated below.

The likelihood a multivariable model has significantly over fitted the data can be assessed using *m* the 'limiting sample size', in this case, the number of recurrences or documented disease related deaths. Studies in which models have been validated on independent datasets have shown that in many situations a fitted regression model is likely to be reliable when the number of predictors, *p* is less than m/10 [11]. The 10:1 rule suggests the potential to analyze up to 12 degrees of freedom.

The relationships between the stage II colon cancer genomic classifier and standard prognostic variables: age at resection; sex; depth of tumor invasion, tumor grade; tumor location, tumor stage, number of nodes examined, and MMR will be presented. The estimated effects (hazard ratio and 95% confidence intervals) will be reported from an analysis including the stage II colon cancer genomic classifier and all other variables in the model. In accordance with REMARK guidelines, the standard prognostic variables and the stage II colon cancer genomic classifier estimates will be reported regardless of significance. The Cox proportional hazards or parametric survival model assumptions will be assessed graphically or by means of a statistical test.

Secondary Exploratory Analysis

As a secondary exploratory analysis, a multivariable model including the genomic classifier will be fit to these data. Stepwise multivariable survival regression will be applied with backward selection. Only factors that show statistically significant association with recurrence with $p<0.01$, based on the likelihood ratio test after adjustment for other factors in the model, will be included in the final model. Internal validation with bootstrap will be applied to validate the modeling process and the stability of the final model. [12]

All microarray data generated from the CALGB 9581 samples including the source data and subsequent analyses will be transferred to CALGB in a mutually agreeable format. The CALGB statisticians will receive the classifier prediction from Almac and will conduct all statistical analyses.

5.0 ADMINISTRATIVE SUMMARY

Selection of Study Sample

Five hundred fourteen (514) patients were randomly selected from the available 901 patients meeting the entry criteria. The total number of eligible patients with available samples (n=902) dropped by one patient from the original submission based on an inventory update from the PCO. Seventy-five (75) patients meeting the recurrence endpoint were included in the random component of the study sample. Under the case-cohort design, the remaining 49 patients meeting the recurrence endpoint were also part of the study sample. A total of 563 patients were identified for potential inclusion in the study. 549 patient samples were provided to Almac. There were 14 blocks that were either missing or unable to be cut because they were insufficient.

Microarray Analysis Results

Of the 549 samples provided by the Alliance sufficient total RNA was extracted to allow for gene expression profiling of 520 samples. Among the 29 samples that were not profiled by gene expression, 6 were unable to be extracted due to lack of tissue availability and 23 due to insufficient RNA. Of these 520 samples, 510 samples were hybridized to a microarray generating raw data (CEL files). Ten (10) samples were not hybridized to a microarray due to insufficient cDNA. Of these 510 samples with microarray data generated, 488 passed ColDx QC metrics (a total of 22 QC failures).

An Excel file containing the following data fields was provided to the CALGB for the 510 samples that generated microarray data, including the 22 QC failures: CALGB Pathology Coordinating Office Number; PTI ID assigned to each sample by Almac (for correlation to raw CEL files); ColDx signature score; dichotomized clinical call (Low Risk, High Risk, QC Failure). The subgroup of 488 samples passing ColDx QC metrics comprised 449 samples including 69 events in the randomized sample and 39 events outside of the randomized sample.

Statistical Analysis of ColDx data

The ColDx data provided by Almac were merged with the clinical data from CALGB 9581 and analyzed for association with recurrence-free interval (RFI) in R using the "survival" package by T. Therneau (Self-Prentice method). [13] The resulting hazard ratio estimate (hr) for the ColDx score categorized at 0.4377 (high risk versus low risk) was 1.189 with 95% CI of (0.69, 2.02). A hazard

ratio of this small a magnitude was not expected by Almac given the previous validation of the ColDx score. The reported validation hazard ratio for RFI was 2.53 (95% CI of (1.54, 4.15). [14]

To verify that the current results were correct the laboratory and clinical data merge, RFI endpoints, and statistical analyses were reviewed. Patient level data were provided to Almac for exploration of potential issues related to quality control, macro-dissection, and other laboratory procedures. In this process, a strong batch effect related to a specific reagent lot number (1201176-B) was noted. The reagent failure coincided with the reagent lot responsible for poor cDNA yield and the subsequent halt to profiling of CALGB samples in March 2012 (hr, 0.84; 95% CI (0.35, 1.97); n=363), while the remaining samples profiled with a good reagent lot exhibited the expected test performance (hr, 3.05; 95% CI (1.27, 7.32); n=147). It was concluded that a lower than expected hr was observed due to reagent failure in the first set of CALGB samples profiled; the reagent lot batch effect was observed at the global gene expression level (PCA) and this effect persisted through to the signature score. Other technical factors were also evaluated and eliminated as causes.

To salvage the validation and maintain the integrity of the study the following steps will be followed. This process is approved by CTEP.

1. All residual samples from all batches will be relabeled by the Alliance Pathology Core to re-establish blinding with assay results.
2. All computer code and raw data from the original assays will be sent to the Alliance Statistics and Data Center (SDC) where the code on the data will be run and the full set of initial predictor results verified.
3. The blinded, relabeled samples will be available to Almac for reanalysis.
4. A QC metric will be put in place that will qualify any given lot of the amplification reagent used in the Col-Dx assay. (This reagent lot qualification test, RLQ test, will not be run on clinical samples as a sample by sample QC metric.)

   The reagent failure occurs at just one key step in the process, the reverse transcription and amplification of RNA into cDNA. There is no evidence to suggest that the RNA extraction step is contributed to the failure observed. Evaluation of historical data generated using multiple lots of this reagent and >120 replicates of a well-characterized FFPE-embedded cell line (colorectal cancer cell line HCT116) indicates that the failure observed in the prior analysis has not been experienced previously. That is, the failure observed is not part of the natural variation of the reagent performance. These data were used in the development of the RLQ test.

   The RLQ test will be implemented by evaluating the reagent lot using multiple runs of the FFPE HCT116 cell line 12-reaction kit provided by the reagent manufacturer profiled on the microarray. The output of the RLQ test is a similarity score based on a subset of genes on the array, which will allow identification of reagent lots that are not suitable for use. Any reagent lot being used to run the Col-Dx assay must pass this metric in order to be released to the lab for clinical sample testing.

5. Almac will send the raw data for the re-done assays to the Alliance SDC where the predictor on the new raw data will be run and the new prediction results will be obtained. Almac will also analyze these data to confirm the Alliance results.

6. A manuscript will be written and submitted for publication within 6 months of completion of the assays. The paper will describe the entire sequence of events; a report of the initial failed validation will be provided.
7. CTEP must review and approve the manuscript before its submission to a journal.

6.0 STATISTICAL CONSIDERATIONS (REVISED)

Residual samples from all batches will be relabeled by the Alliance Pathology Coordinating Office to re-establish blinding with assay results and the Alliance Statistics and Data Center (SDC) will provide Almac a randomization schema by which the re-labeled samples will be re-assayed. Almac will send the raw data for the re-done assays to the Alliance SDC where the predictor on the new raw data will be run and the new prediction results will be obtained.

A case-cohort design will be used to compare recurrence-free interval (RFI) between high risk and low risk patients as determined by the gene expression signature. RFI is measured from study entry until recurrence of primary cancer or death due to primary cancer. Power computations are based on the current randomized subcohort comprising 449 patients with 69 patients meeting the RFI endpoint. Random selection was stratified on treatment (observation; edrecolomab), though survival in the parent study (C9581) did not differ between treatment groups. Thirty-nine (39) patients meeting the RFI endpoint are outside the random subcohort. The assumed final size of the analysis data set is, thus, expected to be 488 patients, including 108 events. With 449 of 901 patients sampled in the random sub-cohort (a sampling fraction of 0.4983) and the following assumptions: the proportion of samples "marker positive,"p1, is 0.20; the proportion of events in the full cohort, pD, is 0.1199 (108/901); 2-sided α=0.05, a hazard ratio of 2.05 can be detected with approximately 80% power. This hazard ratio is lower than that previously observed. [14] Power estimates were determined based on the methods proposed by Cai and Zeng [7] Data analyses will be conducted in R using the 'survival' package by T. Therneau (Self-Prentice method).

7.0 REFERENCES

1. Jemal A, Siegel R, Ward E, et al: Cancer Statistics, 2007. CA Cancer J Clin 57:43-66, 2007.
2. Kuebler JP, EWieand HS, O'Connell MJ, et al: Oxaliplatin combined with weekly bolus fluorouracil and leucovorin as surgical adjuvant chemotherapy for stage II and III colon cancer: Results from NSABP C-07. J Clin Oncol April 30 EPub 2007.
3. Wolmark N, Wieand S, Kuebler JP, et al: A phase III trial comparing FULV to FULV+oxaliplatin in state II or III carcinoma of the colon: Results of NSABP protocol C-07 [abstract]. J Clin Oncol 23, 2005.
4. Prentice R.L., A case-cohort design for epidemiologic cohort studies and disease prevention trials. Biometrike: 73(1), 1986 pgs 1-11.
5. Colacchio T, Niedzwiecki D, Compton C. Phase III trial of adjuvant immunotherapy with MOAb 17-1A following resection for stage II adenocarcinoma of the colon (CALGB 9581). *Proc ASCO 2004; Abstract #3522.*
6. Niedzwiecki D, Bertagnolli MM, Warren RS, et. al: Documenting the natural history of patients with resected stage II adenocarcinoma of the colon after randomization of adjuvant treatment with edrecolomab or observation (CALGB 9581), J. Clin. Oncol. 29(23): 3146-52, 2011, PMID:21747085.
7. Cai J and Zeng D. "Sample size and power calculation for case-cohort studies." Biometrics 60(4): 1015-1024, 2004.
8. Chen K, Lo S. Case-cohort and case-control analysis with Cox's model. Biometrika. 1999;86(4):755-764.
9. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. Springer; 2001.
10. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM: Reporting recommendations for tumor marker prognostic studies (REMARK). J National Cancer Institute. 97 (16): 1180 – 1184, 2005.
11. Riley RD, Sauerbrei W, Altman DG. Prognostic markers in cancer: the evolution of evidence from single studies to meta-analysis, and beyond. British Journal of Cancer, Apr 21; 100(8):1219-29, 2009.
12. Harrell, Jr. FE, Springer-Science. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression and Survival Analysis. 2001.
13. Therneau T. R Package 'survival' URL http://r-forge.r-project.org. Repository CRAN. April 25, 2012.
14. Kennedy RD, Bylesjo M, Kerr P, et. al. Development and independent validation of a prognostic assay for stage II colon cancer using formalin-fixed paraffin-embedded tissue. JCO 29(35): 4620-26, 2011.