



Bayesian Machine Learning on CALGB/SWOG 80405 (Alliance) and PEAK Data Identifies Heterogeneous Landscape of Clinical Predictors of Overall Survival (OS) in Different Populations of Metastatic Colorectal Cancer (mCRC)

Rahul K Das^{1*}, Fang-Shu Ou², Cecilia Washburn¹, Federico Innocenti⁴, Andrew B. Nixon⁵, Heinz-Josef Lenz⁶, Charles Blanke⁷, Donna Niedzwiecki⁸, Iya Khalil¹, Brian D. Harms¹, Alan P. Venook⁹

¹GNS Healthcare, Cambridge, MA; ²Alliance Statistics and Data Center, Mayo Clinic, Rochester, MN; ³Dana-Farber/Partners CancerCare, Boston, MA; ⁴University of North Carolina at Chapel Hill, Chapel Hill, NC; ⁵Duke Cancer Institute, Duke University Medical Center, Durham, NC; ⁶University of Southern California, Los Angeles, CA; ⁷SWOG Group Chair's Office and Knight Cancer Institute, Oregon Health & Science University, Portland, OR; ⁸Department of Biostatistics and Bioinformatics, Duke University, Durham, NC; ⁹University of California at San Francisco, San Francisco, CA.

OBJECTIVES

- Identification of prognostic factors specific for a metastatic colorectal cancer (mCRC) patient subpopulation is key for making personalized treatment decisions and better risk stratification at initiation of treatment.
- Alliance CALGB/SWOG 80405 is a phase III clinical trial w/ randomly assigned cetuximab (cet) /bevacizumab (bev) whereas PEAK is a phase II clinical trial w/ randomly assigned panitumumab (pan) /bev.
- Bayesian machine learning approach was applied on these datasets to discover gender and primary side (1^o) specific predictors of OS and PFS in mCRC and to identify patient subpopulations with better response from a given treatment.

STUDY POPULATIONS & DATA

- 2110 patients from CALGB/SWOG 80405 and 228 patients from PEAK study
- 28 clinical variables as potential predictors
- Sample sizes for patient subpopulations
 - Cetuximab: 831; Bevacizumab: 928; Panitumumab: 116; Cetuximab + Bevacizumab: 463
 - Male: 1373; Female: 965
 - Left: 1482; Right: 856
- Outcome measures: Overall Survival (OS) & Progression Free Survival (PFS)

METHODS

- GNS Healthcare's proprietary Bayesian machine learning platform, Reverse Engineering and Forward Simulation (REFSTM) was used to build predictive models.
- Selection of a single model underestimates prediction error, thus REFS learns an ensemble of the most probable models (N=128) given the data.
- Ensemble constructed via Monte Carlo sampling of the posterior model landscape.
- Model additions/ subtractions scored based on a maximum entropy structural prior with complexity also penalized by the Bayesian Information Criterion.
- Linear, additive, quadratic, and cubic terms allowed in order to accommodate non-linear effects and sub-populations.
- Confidence of a given relationship X->Y determined by frequency among ensemble.

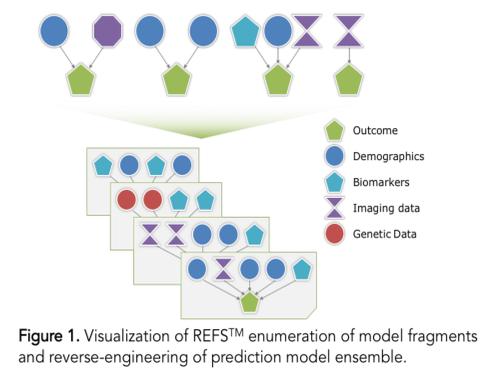


Figure 1. Visualization of REFSTM enumeration of model fragments and reverse-engineering of prediction model ensemble.

- 9 independent ensembles of predictive models were built for different patient subgroups
- Model performance was evaluated using Concordance Index (CI)
- p values of the interaction term were obtained from univariate Cox Proportional Hazard model

RESULTS

REFS MODELS UNCOVERED HETEROGENEOUS LANDSCAPE OF CLINICAL PREDICTORS OF OS AND PFS OUTCOMES IN mCRC PATIENTS

Biomarkers	Full (n=2338)		Left (n=1482)		Right (n=856)		Female (n=965)		Male (n=1373)		Cet (n=831)		Bev (n=928)		Pan (n=116)		Cet + Bev (n=463)		
	OS	PFS	OS	PFS	OS	PFS	OS	PFS	OS	PFS	OS	PFS	OS	PFS	OS	PFS	OS	PFS	
AST	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Intra abdominal metastasis	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Lung metastasis	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
1 ^o side (Rvs.L)	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
HGB	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Urine protein level	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
ALB	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
NEUT (%)	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●

↑ Negative prognostic
↑ Positive prognostic
● Edge frequency > 50%
● Edge frequency 20%-50%

Table 1: List of ensemble edge frequencies of different predictors of OS and PFS from REFS models built on different patient cohorts. The edge frequency quantifies importance of a variable as predictor of the outcome. e.g. if a variable appears as predictor of the outcome in all 128 models, it will have 100% edge frequency and is very likely a strong predictor.

KAPLAN MEIER SURVIVAL CURVES FOR PATIENT SUBGROUP SPECIFIC PROGNOSTIC FACTORS OF OS AS IDENTIFIED BY REFS MODELS

Figure 2: KRAS status and Hemoglobin level: prognostic factors of OS in only left tumors

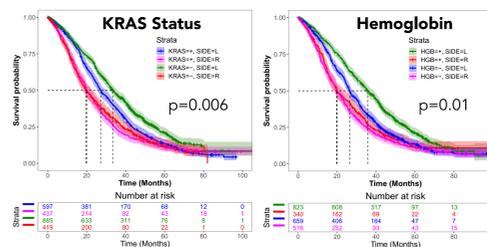


Figure 3: Creatinine level and intra-abdominal mets: prognostic factors of OS in only females

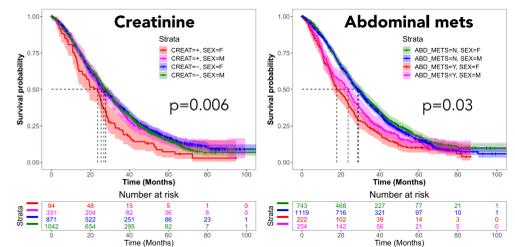


Figure 4: Neutrophil (%) level: prognostic factor of OS for females with lower ALB level (log rank p = 2.3 x 10⁻⁶ in females, 0.01 in males)

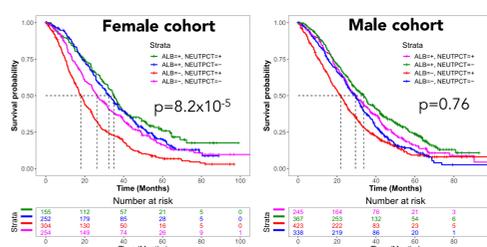
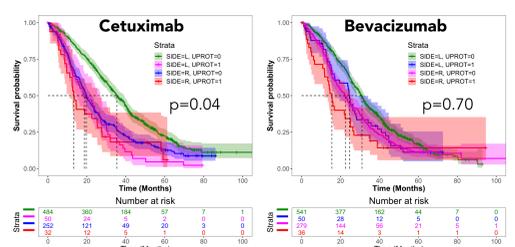


Figure 5: Better efficacy for left- than right-sided patients from Cet treatment only when urine protein level is trace



MODEL PERFORMANCE INDICATIVE OF IMPORTANCE OF MOLECULAR BIOMARKERS

Models	OS		PFS	
	CI (In sample training recovery)	CI (5-fold cross validation)	CI (In sample training recovery)	CI (5-fold cross validation)
Full	0.63	0.62,0.62,0.61,0.64,0.61	0.56	0.55,0.53,0.55,0.56,0.55
Female	0.65	0.64,0.63,0.65,0.61,0.58	0.57	0.57,0.56,0.56,0.55,0.54
Male	0.61	0.61,0.56,0.58,0.59,0.58	0.50	0.51,0.49,0.50,0.50,0.50
Left-sided	0.63	0.63,0.57,0.62,0.62,0.62	0.55	0.53,0.51,0.53,0.53,0.52
Right-sided	0.59	0.61,0.56,0.56,0.60,0.56	0.53	0.51,0.51,0.51,0.50,0.51
Cetuximab	0.64	0.65,0.60,0.62,0.62,0.60	0.58	0.54,0.55,0.57,0.54,0.53
Bevacizumab	0.60	0.59,0.62,0.60,0.63,0.54	0.54	0.54,0.54,0.55,0.54,0.48
Panitumumab	0.61	0.59,0.56,0.56,0.58,0.53	0.56	0.55,0.52,0.51,0.56,0.50
Cet + Bev	0.59	0.61,0.58,0.55,0.56,0.59	0.52	0.50,0.50,0.49,0.51,0.51

Table 2: List of REFS model performance quantified by concordance index (CI)

- Performance of OS models was better than the PFS ones
- Models for females, left tumors, and Cetuximab arm performed better than other models
- Performance was inferior to previous REFS models^{1,2} built from CALGB molecular data, **suggestive of importance of molecular variables as prognostic factors of OS and PFS**

Cetuximab model (in sample concordance index=0.65; 95% confidence interval=0.63-0.67) had comparable out-of-sample predictive performance in Panitumumab (0.68; 0.62-0.74) and Bevacizumab (0.66; 0.60-0.72) patient cohorts

CONCLUSIONS

- Bayesian machine learning was used to build ensembles of predictive models for OS and PFS outcomes for different patient mCRC populations using clinical variables from CALGB 80405 and PEAK trial data
- Landscape of the predictors across different patient mCRC populations and outcomes was heterogeneous
 - AST level, 1^o side, intra-abdominal mets status were ubiquitous predictors
- Side specific predictors of OS, such as KRAS status and levels of hemoglobin were identified
- Gender specific predictors, such as creatinine level (OS and PFS), intra-abdominal mets status (OS), and interaction of albumin and neutrophil (%) levels (OS and PFS) were identified
- Patients subpopulations who benefitted from different treatment arms were identified
 - Urine protein level was predictive of better efficacy (OS and PFS) from Cetuximab treatment in left-sided tumors
- Sample size of the Panitumumab treatment arm was too low to identify predictors and interactions with desired significance level
- The comparable but modest model performance across different treatment arms and studies demonstrates importance of molecular variables, unavailable for this study, for better differentiation of patient response to treatments.

REFERENCES

1. Das et al., Annals of Oncology (2018) 29 (suppl_8): viii150-viii204. 10.1093/annonc/mdy281
2. Das et al., Journal of Clinical Oncology 2018 36:15_suppl, 3570-3570

SUPPORT: U10CA180821, U10CA180882, Eli Lilly and Company, Genentech, Pfizer, Amgen; <https://acknowledgments.alliancefound.org>



*CORRESPONDENCE: rdas@gnshealthcare.com